

A discipline buried by success

Chemistry suffers from multiple image problems. Chemists working at its boundaries should acknowledge and celebrate their roots, while those at its core have much to celebrate too.

Chemistry's goal of understanding the form and function of molecules and studying how they interact is at the very heart of scientific endeavour. Yet its versatility comes at a price: recognition — or, more accurately, lack of it.

Just what is chemistry? The subject is no longer restricted to its academic organic, inorganic and physical divisions. In addition to well-known territories such as catalysis, organic synthesis, polymers and materials science, chemistry has extended to other less familiar areas — the use of small molecules in 'chemical genetics', screening drug candidates with combinatorial techniques and developing new analytical techniques that touch virtually every other part of science. The subject's sheer diversity — the American Chemical Society alone has over 30 divisions — makes generalizations about it impossible.

Perhaps such a definition is unnecessary. But lack of an accurate and identifiable chemistry 'brand' means that the discipline is easily misunderstood, and those working in it are frequently underappreciated. This is the basis of the discipline's infamous image problem. Chemists have allowed those from outside the field to characterize it — to define what chemistry is and what it is not.

To the public, chemical science is too often synonymous with the industry with which it shares its name. So chemistry means belching chimneys and poisoned rivers, not life-saving medicines and space-age materials. To other researchers, policy-makers and, crucially, young scientists, it is sometimes seen as a mature discipline with its most important and stimulating work behind it. Chemists, on the other hand, tell a different story. They speak excitedly about the promise of molecular electronics, the challenges presented by the need for sustainable energy and the opportunities for bespoke pharmaceuticals presented by the human genome sequence.

Analysis in demand

The subject's reputation is not a trivial issue. If chemistry is to achieve half as much in this century as it did in the last, it must fight to attract the brightest young scientists. And to do that, image is important. Furthermore, basic research in both academic and industrial chemistry needs a shot in the arm. Basic chemistry research in companies is suffering as research budgets are tightened. And basic chemistry research in universities is being overshadowed by the continued growth in support for biology. As one chemist laments: "biology is just a lot more attractive to people now."

But chemists and chemistry have never been more vital to science and society than now. As the borders between scientific disciplines blur (a process that will only continue), fundamental chemistry skills such as synthesis and analysis will be crucial for the interdisciplinary subjects that emerge. Indeed, chemists are already flocking in increasing numbers to the edges of their discipline to collaborate with biologists, physicists, engineers and computer scientists.

Research at these fringes is dynamic, fascinating and of great potential benefit. Chemists are vital to projects such as the use of medical research imaging to probe development and the study of molecular recognition and binding using synthetic cell membranes. Chemical analysis of protein folding and techniques based on mass spectroscopy are needed to capitalize on genomics, and it is chemical

self-assembly techniques that are driving the nanotechnology bandwagon. This eye-catching interdisciplinary research offers chemistry a chance to polish its image. But first, chemists must learn to celebrate and promote their contributions.

As we discuss on page 408, one reason for chemistry's tarnished image is that rival fields have claimed many of its brightest moments. The reasons for this are subtle and numerous, but, as George Whitesides at Harvard University suggests, modesty could play a role. If that's true, chemists should stop hiding their Bunsen burners under bushels.

Championship gaps

Chemical societies are evolving to acknowledge the changing face of the subject they represent. In Germany, 2003 will be the year of chemistry, as the German Chemical Society publicly celebrates the subject's diversity. In Britain, the Royal Society of Chemistry is relaxing its membership criteria so that those doing chemistry, but who lack a classical chemistry background, can join. The American Chemical Society is trying to stamp chemistry's brand on emerging topics such as nanotechnology and proteomics through new divisions, journals and conferences.

These are welcome steps in the right direction. But if chemistry is to lay stronger claims to its future achievements, more of the front-line chemists streaming across the discipline's borders into attention-grabbing multidisciplinary research must make their voices heard. They should proclaim their roles to colleagues and try to ensure that chemical contributions are made known to the media.

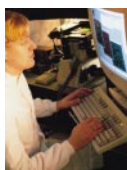
This may be difficult — the biological aspects of such projects are often easier for non-experts to understand. But the 1,000 chemists attending the recent meeting of the American Chemical Society in San Diego produced just a handful of press releases between them. If chemists want a better image they must lose their modesty.

Perhaps then, more of the chemical-sciences limelight will reflect back onto 'core' chemistry topics that also deserve greater attention. There are many. A recent poll of US chemists by the head of chemistry at the Massachusetts Institute of Technology, Stephen Lippard, produced a 'wish-list' of more than 20 exciting areas in basic chemistry that are ripe for discovery (see *Nature* **406**, 807–808; 2000). These include self-replicating molecules, the search for reagents and pathways capable of breaking inert chemical bonds, designing customized catalysts for some reactions and eliminating the solvents from others. Such developments, and an improved yield, specificity and versatility of chemical reactions by greater control over the molecules involved, are, as Lippard says, driving a "quiet revolution" in chemistry.

But here, too, chemists should help their societies tell undergraduates, schools and the media what these breakthroughs could lead to. Tellingly, Lippard stops short of identifying potential applications of his wish-list to surrounding fields, and therefore their possible implications for the world at large. "We are confident our colleagues in those areas can make these connections," he says.

That is surely a missed opportunity. If more chemists established these connections themselves, and talked up the potential benefits, their contribution would not be so easily overlooked. Revolutions, after all, should be noisy. ■

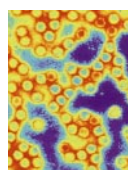




Personal service
Second genome sequencing revolution in sight
p402



Wipeout?
Lack of funds could halt eradication of rinderpest
p403



Unwanted return
Polio defies medics and attempts a come back in Europe
p405



Caught out
DNA database snares illegal salmon
p407

Russia fuels fury with scheme for importing nuclear waste

Quirin Schiermeier

Russia is planning to import 20,000 tonnes of spent nuclear fuel for storage and possible reprocessing in Siberia, despite international opposition from environmental groups.

The nuclear ministry, Minatom, hopes the scheme will generate US\$20 billion over the next 12 years. The highly radioactive fuel will be stored, and perhaps treated, at two facilities near Majak and Krasnojarsk.

The importation of spent fuel is currently forbidden under Russian environmental laws. But the Russian State parliament looks set to approve a bill, which is backed by President Vladimir Putin, that would lift the ban.

Minatom has already approached possible customers in Taiwan, South Korea, Switzerland and Japan. And a Minatom delegation lobbied hard for the plan at an international meeting on nuclear-waste technology in Dresden last week.

But Minatom can expect powerful opposition to the scheme. The US government controls 90% of the world's spent fuel, and is almost certain to block the transfer of fuel from the United States to Russia. Plutonium is extracted during reprocessing, and the United States is concerned about the possible proliferation of weapons-grade plutonium.

The remaining 10% of spent fuel — around 20,000 tonnes — is stored by coun-



Left cold: activists protest at Minatom's plans to reprocess spent nuclear fuel in Siberia.

tries in eastern Europe and Asia, which are less likely to be able to afford the price tag for storage envisaged by Minatom of \$1 million per tonne.

"The Minatom proposal is a guaranteed non-starter," says Tom Cochran, director of the Nuclear Program of the Natural Resources Defense Council, a US-based environmental organization.

Despite this, an alternative scheme that would allow Russia to store but not reprocess spent fuel is being met with relative disinter-

est from Russian officials. The plan was developed by the Non-Proliferation Trust (NPT), a Delaware-based charitable corporation that promotes non-proliferation of nuclear material, as well as environmental and humanitarian initiatives in Russia.

The NPT aims to raise \$15 billion by building and operating an interim storage facility in Russia, with 10,000-tonne capacity, for spent fuel from sources outside the United States. Profits would be used for clean-up and charity projects in Russia, for

Stanford agrees to settle gender discrimination case

Rex Dalton, San Diego

A computer scientist has won a major settlement against Stanford University in her lawsuit for gender discrimination and retaliatory firing.

The settlement between Colleen Crangle and her former employer was announced last week in Palo Alto, California, with neither side revealing the terms of the agreement. But *Nature* has learnt that Stanford paid \$950,000 to close the books on the case, in which the university was appealing after losing a jury trial last year (*Nature* 404, 536; 2000).

"This allows the parties to avoid further

protracted appeals and litigation and to focus their efforts on the important work before them," says Debra Zumwalt, Stanford's general counsel.

This is at least the third settlement in recent years of gender-discrimination lawsuits brought by woman researchers at Stanford. Top officials at the university have acknowledged that barriers still hamper women from advancing in science and engineering. The US Department of Labor has been investigating allegations of gender discrimination at Stanford since 1999.

"These women are sending a strong message that sex discrimination will not be

tolerated," says Sylvia Newman, president of the American Association of University Women Legal Advocacy Fund, which provided seed money for Crangle's case.

The settlement comes as reports have found disparities in how women researchers are treated and paid at some US universities. Last week, the University of Arizona's College of Medicine in Tucson released a study showing that the average salary of female researchers was 11% lower than that of males. Women faculty members, including physicians and basic research scientists, were paid \$13,000 less than their male colleagues, according to the study. ■

improving nuclear safety technology, and for building a deep geological repository in Siberia to dispose of the nuclear waste.

The project, which requires approval from both the US and Russian governments, is conditional on Russia agreeing not to engage in commercial reprocessing activities for 30 years.

"Such a scheme would be beneficial to Russia and to the rest of the world," says Cochran. But Minatom seems unwilling to make the concessions required, he says.

Environmentalists are unhappy with both plans. Majak and Krasnojarsk are among the most polluted spots on Earth, says Tobias Münchmeyer, a nuclear campaigner at Greenpeace International. "It would be cynical to use these ecological disaster areas as dumping grounds for foreign nuclear waste."

Vladimir Sliviyak, co-chairman of Ecodefense, a network of environmental groups in Russia, says that Minatom would find ways to ignore or alter any international agreement, he says. "Whoever thinks that Minatom can be influenced by foreign powers is either money-driven or naive," he says. "If waste gets into Russia you can bet that it will be reprocessed and that plutonium will be separated."

GM cows face slaughter in multiple sclerosis experiment

Bob Brockie, Wellington, New Zealand

The New Zealand High Court has granted a stay of execution on its ruling to slaughter a small herd of cows pregnant with genetically modified calves.

The unborn calves are part of a programme aimed at developing a treatment for multiple sclerosis. Researchers at life-sciences company AgResearch have inserted the gene for human myelin basic protein into the



Errors in law could halt studies using GM cows.

embryos in the hope of producing high levels of the protein in the cattle's milk. They hope to use the protein to develop therapies for multiple sclerosis.

The Environmental Risk Management Authority (ERMA), the statutory body that vets any genetic engineering in New Zealand, approved the experiments last year. But activists and Maoris challenged the approval and the High Court ruled, for a second time, that the experiment must stop.

The court did not question the science behind the research but ruled that ERMA made errors of law in its approval process. ERMA has a week to reconsider its approval and to present details of its decision to the court, or the experiment must stop. If ERMA fails, the cattle will be slaughtered on 28 May.

The Green Party, which opposes any form of genetic modification, is calling for the immediate slaughter of the cows. The head of the research programme, Phil L'Hullier, frustrated at the delays, recently resigned to pursue similar work in Australia.

Individual genomes targeted in sequencing revolution

David Adam, London

A Cambridge biotech start-up company is developing new technology that it claims could allow an individual's genome to be sequenced in just a few days — at a fraction of the cost of existing technology.

Strong claims indeed — today's highest-throughput sequencing labs currently take months to sequence a mammalian genome completely. But Solexa, which opened its new laboratories earlier this month, may have the scientific muscle to back up such claims. Three members of the company's scientific advisory board hold top positions at the nearby Sanger Centre, which decoded around a third of the human genome.

Solexa hopes its technology will help to identify the complex associations between disease and the sites of genetic variability, called single nucleotide polymorphisms (SNPs), which are scattered through the genome. SNPs are time-consuming and expensive to locate in just one individual, never mind in the tens of thousands of patients and controls that would be needed for large-scale studies of such associations.

"It's a numbers game and nothing around at the moment gets anywhere close to being powerful enough to do those kind of genome-wide association studies," says Nick McCooke, Solexa's president

and managing director.

Solexa is also some way off that goal, McCooke admits. But he claims that within two years his company could produce a working prototype that would have "the sequencing capacity of 20 Sanger Centres". He says comparing the full sequence of a genome is the best way to sift through individuals to identify the crucial genetic differences, including SNPs.

At the heart of the new technology, which is still to be finalized and scaled up, are what the company terms 'single-molecule arrays'. The single-stranded DNA samples to be sequenced, comprising strings of the four nucleotides adenosine (A), cytosine (C), guanine (G) and thymidine (T), are first immobilized onto a surface through a synthetic sequence of 'primer' DNA. Solexa says it will produce chips that can hold up to 100 million immobilized samples.

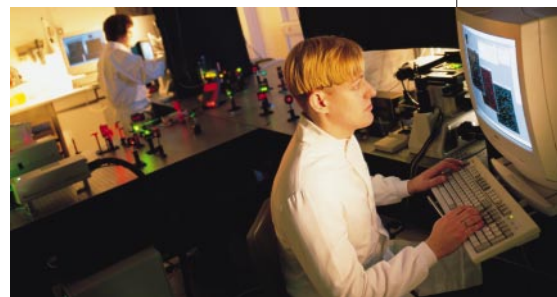
The surface is sequentially washed with solutions containing a single type of nucleotide — tagged with a fluorescent marker — and a DNA polymerase enzyme. The enzyme tries to build the complementary strand — A binding with T, and G binding with C — starting at the primer sequence. After each wash, a laser scanner identifies which strands have gained complementary nucleotides. The fluorescent

tags are then chemically removed and the process is repeated. The complementary strands thus build up one nucleotide at a time, so analysis is simple.

In this way, each of the 100 million immobilized DNA samples acts like a sequencing lane in a conventional machine. And because the technique requires no initial amplification of the samples and very small quantities of reagents, McCooke says it will be far cheaper than existing technology.

Solexa is not the only company working on such sequencing techniques. Amersham Pharmacia Biotech, for example, says it is developing a single-molecule array technique that is "similar in scope".

♦ www.solexa.co.uk



Getting personal: Solexa's new technique could sequence an individual's genome in a few days.

More funding needed to wipe out rinderpest

Sally Goodman

While Britain and Uruguay continue to battle against foot-and-mouth disease, another devastating livestock disease, rinderpest, could be on the point of total eradication. It would be only the second disease in history, after smallpox, to have been wiped out.

But researchers warn that if funding for the final stages of the eradication programme remain unsecured, the efforts so far made to eliminate the disease will be wasted.

Rinderpest is a debilitating and often fatal viral disease that primarily affects cattle, buffaloes, yaks and wild ungulates. It has had catastrophic effects on agriculture in Africa and Asia in the recent past. But an extended vaccination and international control programme has confined the disease to a few isolated reservoirs in Pakistan and the war-torn areas of Somalia and southern Sudan, where control has been more difficult.

According to Peter Roeder of the Global Rinderpest Eradication Programme, which is coordinated by the UN Food and Agriculture Organization, the world has never been this close to the eradication of rinderpest. "But there is a real danger of complacency setting in," he says, "and a real risk of epidemic or even pandemic recurrence if control efforts falter."

Where it still exists, rinderpest occurs in a mild form and could persist unnoticed should it reach Europe or the United States through international livestock trade.

Roeder estimates that a further US\$15 million are needed to fund a final eradication effort over the next five years. This should focus on disclosing and eliminating any remaining reservoirs of the virus once vaccination has ceased, he says, and absence of disease should be verified through serological surveillance. He also recommends that a global emergency programme be put in place to deal quickly with any resurgence.

Most research into rinderpest is funded by the European Union through the Pan-African Programme for the Control of Epizootics, which will end in 2004. The UK Department for International Development and the USAID programme also support fieldwork in Somalia and Sudan. Laboratory research is mainly focused on new-generation vaccines, which could be used to 'ring-fence' vaccination, and on the development of more sensitive and specific diagnostic kits to differentiate between rinderpest and another closely related disease, peste des petits ruminants.

The Pirbright laboratory of the Institute of Animal Health (IAH) in Surrey, UK, houses the World Reference Laboratory for rinderpest. Staffed by only three people, it has recently developed a diagnostic kit that uses the same 'dipstick' technology used in



On its last legs: the livestock disease rinderpest could soon be eradicated if funding is maintained.

home pregnancy tests. The kit, which gives results within 10 minutes, is already being used in Pakistan. John Anderson, who heads the development team, says that such tests will be vital in the final surveillance stage of the eradication process.

The IAH is also developing marker vaccines that will differentiate between infected and vaccinated cattle. And a national agriculture laboratory in Montpellier, France, has begun work to develop an oral vaccine in pellet form.

But funding is most urgently needed to develop capacity to fight the disease in the field. Affected countries must have adequate local resources and expertise to take part in the final stages of the eradication programme, says Manzoor Hussain, an animal health officer in Pakistan. "The main problem is that we lack efficient reporting systems," he says.

Meanwhile, Anderson has had to stop his own work on rinderpest to concentrate on foot-and-mouth testing in Britain. ■

Plans for GM livestock fail the poor

David Adam, London

Research to develop genetically modified (GM) animals is focusing on producing cheaper food for the West, and risks ignoring the technology's potential to help agriculture in the developing world, the Royal Society warned this week.

A report on the use of GM animals calls for greater investment for work to create animals resistant to foot-and-mouth disease and sleeping sickness. "Creating disease-resistant animals is especially important for the farmers in the developing world and the society recommends that research efforts on this technology are addressed with particular urgency," the report says.

It calls for greater cooperation between universities and biotechnology companies to develop disease-resistant GM animal technology and for "a willingness to share knowledge currently restricted under patent and licensing agreements".

Bob May, the society's president, says big

agricultural businesses have little incentive to address problems unique to countries that cannot afford GM technology. The recommendation "must be taken seriously because GM technologies could have great benefits for developing countries," he says.

The report also recommends a ban on rearing GM fish in marine pens. Approval for commercial production of GM salmon should only be given if the fish are kept in landlocked tanks, it states. This echoes recommendations made by a panel from the Royal Society of Canada earlier this year (see *Nature* 409, 749; 2001). A US company, Aqua Bounty Farms, has developed GM salmon that grow faster than normal, prompting fears that the fish will disrupt natural salmon populations if they escape (see *Nature* 406, 10–12; 2000).

All laboratories working on GM animal technology should have emergency plans for containment, the report adds, in case the animals escape or are released deliberately. ■

Company tells researchers to look to profits

Corie Lok, Washington DC
and David Cyranoski, Tokyo

After a decade of freedom, researchers at the NEC Research Institute (NECI) in Princeton, New Jersey, are now under pressure to conduct research directly relevant to NEC's core business.

When the institute was set up by the electronics and information technology giant in 1989, its blue-skies policy and generously funded laboratories attracted top academic physicists and computer scientists from all over the United States.

But as part of a trend towards more business-oriented research at corporate and university laboratories, NECI changed its mission statement in February to emphasize its shift towards application-based projects.

The research staff of 70 are now receiving feedback from the administrative centres in Tokyo and Princeton about what areas should be prioritized. Changes have already been made to some research programmes — and it is an open secret that some key staff are seeking jobs back in the academic world.

"I've asked the scientists to spend the majority of their time on something with a business connection," says NECI president David Waltz, adding that the institute is still pursuing basic research in potential high-



pay-off areas such as bioinformatics and quantum computing that are 10 to 20 years from commercialization. Research projects will still be proposed and evaluated by NECI's scientists, he says. "NEC will not micromanage the institute." Although this year's budget for the laboratory has been held flat, he says, the institute will be hiring more staff this year.

The change in mission statement was one of a series of moves by the institute and NEC to focus more on research related

to core businesses. Over the past three years, the institute has shifted its emphasis towards computer science and away from the physical sciences, and NEC now offers financial incentives to encourage scientists to carry promising research towards commercialization.

"The changes make for a more uncertain and fragile environment," says William Bialek, a senior NECI scientist who left a professorship at the University of California at Berkeley to join the institute in 1990.

Scientists say it is still too early to see any effects of this change in mission, although some say they are actively "re-examining their career options".

Similar changes have been made at NEC's other research facilities in Japan. Last year the lab in Tsukuba moved its material scientists from basic research into development, says Jun'ichi Sone, the lab's senior manager. It is now concentrating its fundamental research in three areas: nanotechnology, quantum communications and biotechnology.

A decade of research without having to worry about an immediate impact on NEC's profitability was, perhaps, "anomalously good", says Bialek, a neurophysicist. "Maybe now it's becoming a bit more normal." ■

Europe's biotech industry still losing to US, say analysts

Quirin Schiermeier, Munich

The European life-sciences industry has shown a "dramatic improvement" in its financial state, according to a report from a leading business consultant. But despite this, the report says the Atlantic divide in biotechnology is widening.

Integration: Ernst & Young's Eighth Annual Life Sciences Report 2001 says biotechnology companies in Europe last year raised almost 6 billion euros (US\$5.3 billion) on the stock markets, and a record 39 companies went public, offering their shares on the market for the first time.

The United States has fewer biotech companies than Europe, but they are larger and richer. The 1,273 US companies generated revenues of 23.75 billion euros last year, compared with 5 billion euros generated by Europe's 1,570 companies.

For the first time, Germany has taken pole position in European biotech — at least in its number of companies. Germany (330 companies) is followed by the United Kingdom (270), France (180) and Sweden (170). But strongest is the United Kingdom, with its companies contributing almost half of the pan-European revenues in

biotechnology. Many of the young German companies are still growing, and the total number is likely to fall over the coming years through mergers and acquisitions.

Last year, European biotech companies increased their spending on research and development by almost 50%. Net losses, however, increased at almost the same rate.

Public investment in genomics has made biotech a cornerstone of knowledge-based economies in Europe, according to Ernst & Young. "There are some excellent companies with world-class leadership and unbeatable technologies," the report says.

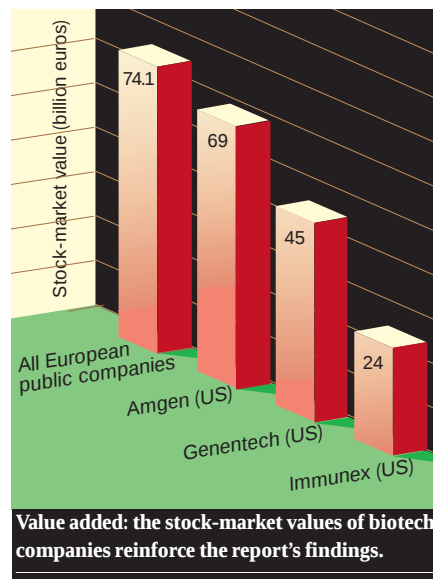
But the consultants warn that the European biotech sector is still moving too slowly compared with its US counterpart. They attribute this primarily to Europe's lack of a true single market and to a "different investors' mentality".

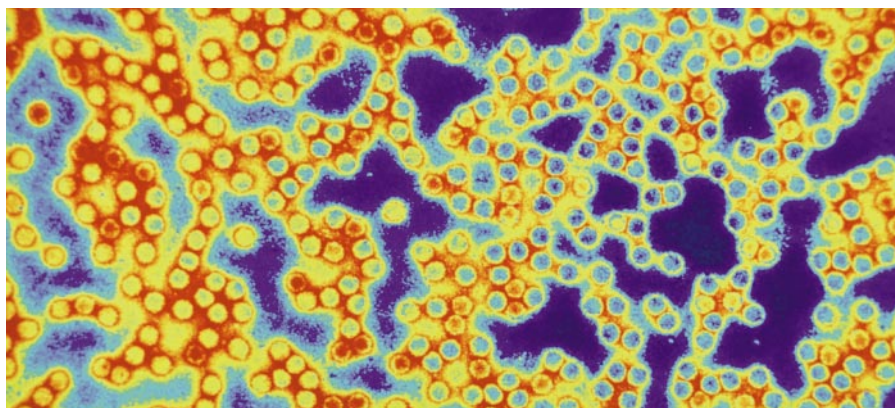
"US investors seem much more willing to provide the hundreds of millions of dollars needed to grow companies aggressively into jumbo-sized packages," says the report. The entire European public life-sciences industry is only slightly larger than the US biotech giant Amgen (see graph).

And size matters, says the report. "If

there are companies in the US trying to solve the problem with five times the resources it's not difficult to imagine who will get there first in most cases." ■

♦ http://www.ey.com/global/gcr.nsf/UK/Ernst_&_Young_Life_Sciences





Fighting back: a wild virus strain from India has given Europe its first polio cases for two years.

Plans to eradicate polio hit by virus outbreak in Bulgaria

Mark Schroppe

Two new cases of poliomyelitis in Bulgaria threaten to delay plans by the World Health Organization (WHO) to eradicate the disease from the planet. The cases raise particular concern because they are caused by a wild virus, rather than by the strain used in vaccinations.

There had been no reported cases of polio in Europe for more than two years. But in March a girl in the Bulgarian seaport of Burgas was paralysed by the disease, and a second case of paralysis was confirmed in the area on 11 May.

Researchers say the outbreak sends a warning to countries around the globe to maintain their guard against polio, despite the rapid decline in its incidence. Bulgaria had not had a case of polio caused by wild virus for almost ten years.

Gene sequences of the strain responsible match a wild strain from India. "It is a matter of real concern whenever you have importation into a disease-free area," says Donald Henderson, an epidemiologist at Johns Hopkins University in Baltimore, Maryland.

The second paralysis case indicates that the disease is spreading rapidly, says George Oblapenko, WHO medical officer for polio eradication, based in Copenhagen. The polio virus only causes paralysis in one person in 100 or more infected, and the symptoms are usually mild and transient — so each case indicates a far greater problem.

Both of the affected children belong to Bulgaria's gypsy population, the Roma, whose transient lifestyle makes effective vaccination difficult. The WHO has already spearheaded a programme to test and vaccinate children in the immediate area of the infections. Oblapenko says preparations are under way for a massive vaccination effort in the coming weeks targeting around

130,000 Roma children as well as hundreds of thousands of other Bulgarian children under seven years of age.

Oblapenko has praised the quick identification of the cases by Bulgarian health workers. Henderson agrees. "Trying to stop polio is not easy," he says, particularly because so few of those infected show signs of infection.

The Bulgarian cases emphasize the need for countries to maintain not only strong vaccination programmes, but also diligent surveillance even in areas that have had no cases for some time, says Oblapenko. "If you do not protect certain groups of a population, imported wild poliovirus will find the hole, and hit," he says. The girl paralysed in March had not been vaccinated because her doctor decided not to expose her to the vaccine because her general health was weak.

The WHO aims to eradicate polio by around 2005, and then to phase out vaccination because it would theoretically be no longer needed. One step in that plan is to certify regions as virus-free, a classification researchers hope Europe may achieve by early 2002. Because the Bulgarian cases involve importation and do not suggest the persistence of an endemic wild strain in Europe, they will not necessarily block certification. But Oblapenko says it may be delayed.

The goal of eradication has been complicated by the discovery in Hispaniola of polio caused by a vaccine virus strain that mutated to regain its disease-causing ability (see *Nature* **409**, 278–280; 2001). Henderson says that such problems, in conjunction with the Bulgarian outbreak, mean that vaccinations must continue.

But Oblapenko says the WHO's goals will eventually be achieved. "We are pretty sure we will stop [polio]," he says. ■

Congress hears plan to boost science at environment agency

Tony Reichhardt, Washington

In a move to address criticisms that science at the US Environmental Protection Agency (EPA) takes a back seat to regulatory politics, a bill to strengthen science at the agency is progressing through the US Congress.

The bill, introduced by Vernon Ehlers (Republican, Michigan), would make environmental research a "central mission" of the EPA and create the post of deputy administrator for science and technology. It was passed unanimously last week by the House Science Committee's environmental subcommittee, which Ehlers chairs.

Ehlers' proposed legislation closely follows advice given in a report on the EPA published by the National Academy of Sciences last autumn. The new deputy administrator would be the agency's chief science adviser, responsible for coordinating the EPA's research portfolio, which at present is overseen by various regulatory offices and field centres. The head of the Office of Research and Development, which sponsors about half of the agency's research, would be designated 'chief scientist' at the EPA and given a five-year term.

The bill now moves to the full science committee for approval. Although it is supported by key legislators, including committee chair Sherwood Boehlert (Republican, New York), the EPA has not endorsed it. Its administrator, Christine Todd Whitman, has said she would prefer to wait until the EPA has completed a review of how science and economics play into its regulatory activities.

Decisions about a new high-level science post are probably secondary to the agency's attempts to get itself elevated to the cabinet. Boehlert recently introduced legislation to that effect, but was forced to withdraw it by the White House, which asked for more time to review its options.

William Glaze, a chemist at the University of North Carolina and chair of the agency's science advisory board, testified in March that the bill would "send a strong signal that the Congress and this administration plan to make science a stronger and more integral part of the way EPA conducts its business."

But the EPA's research budget has remained flat for years. And unless the White House starts listening more closely to the advice of scientists, Glaze says there will be "some scepticism" that a deputy administrator can effect real change. ■

List of scientists' industry ties to be posted on Internet

Washington A US consumer-advocacy group has launched an online database that lists scientists and their links with industry.

The Washington-based Center for Science in the Public Interest (CSPI) developed the database, which includes information on commercial funding and consulting work. Most of the researchers listed are in the fields of medicine, nutrition, environment and toxicology. Information about sources of funding for professional, health and non-profit organizations is also provided.

CSPI used a variety of sources, from medical and scientific journals with disclosure policies, to resumés submitted by scientists when working for government committees. Project director Ronald Collins says it is designed to promote openness and not to query the integrity of scientists.

♦ <http://www.cspinet.org/integrity/database.html>

German space researchers question funding policy

Munich The German government has responded to uncertainty over its funding of space science by announcing a four-year

budget instead of the normal one-year plan. But researchers are still concerned that too much is being spent on international projects.

Space research and technology will receive DM8 billion (US\$3.6 billion) between 2002 and 2005, in line with recent spending of around DM2 billion per year. But only DM350 million will go to German projects, and only DM70 million to national basic research. The rest will be spent on international projects, such as Germany's 41% contribution to Europe's overall funding for the International Space Station (ISS).

Günther Hasinger, scientific director of the Potsdam Astrophysical Institute and director of the Max Planck Institute of Extraterrestrial Physics in Garching, says that German astronomers should play a bigger role in ISS-based research to match the scale of German funding.

Amateur astronomers left in the dark

San Diego A group of astronomers at an historic US observatory has reached an agreement with local state authorities over plans to limit light pollution from a new prison.

Amateur astronomers' club Stellafane, which has been based in Springfield, Vermont, since the 1920s, had complained



Light at the end of the telescope? Astronomers have won a campaign to reduce light pollution.

about plans for the new Southern State Correctional Facility, due to be built this summer at a site four miles from the club. Stellafane argued that the proposed security lighting would have allowed too much light to spill upwards, increasing glare in the sky above the club's telescopes. Vermont officials have now agreed to use 'cut-off' lights that limit the amount of escaping light.

♦ <http://www.stellafane.com>

Sorbonne criticized for 'legitimizing' astrology

Paris French scientists have reacted with outrage to the news that Sorbonne University in Paris has awarded a doctoral degree in sociology to French astrologist Elizabeth Teissier.

Scientists fear that Teissier, who was former president Francois Mitterand's 'astrological adviser', will use the award to bolster her argument that astrology should

be recognized by the Sorbonne as a legitimate area of study. Astrophysicist Jean-Claude Pecker, president of the French Association of Scientific Information, and four French Nobel laureates in science are planning to write individual letters of protest to the education minister, Jack Lang.

A similar controversy has arisen recently in India after authorities gave the go-ahead for astrology departments to be created in universities (see *Nature* 411, 227; 2001).

Concern over research on human subjects

Washington A new independent federal office is needed to oversee protection of human subjects involved in scientific research, a report from the National Bioethics Advisory Commission (NBAC) said last week.

Some privately funded studies, such as those that use subjects from fertility clinics or doctors' offices, are not currently subject to federal regulations. The report argues that all research on humans should be overseen by the new office and that local review panels should include at least 25% non-scientists.

The report adds to growing criticism of the way research involving humans is regulated in the United States (see *Nature* 410, 1017; 2001), but the changes it recommends would require Congressional

action, which some observers believe is unlikely under the new administration. The NBAC will cease to exist in October if, as expected, President George W. Bush refuses to renew its mandate.

► <http://www.bioethics.gov>

French cancer charity mends damaged reputation

Paris France's largest cancer charity, the Association pour la Recherche sur le Cancer (ARC), is attempting to put its turbulent past behind it by announcing a national research network supported by a budget of FFfr100 million (US\$13 million) over three years.

The network will coordinate research into cancer therapies derived from post-genomic molecular biology and will foster interaction between different research groups. The ARC's image has suffered since Jacques Crozemarie, its founder and former chairman, was found guilty in 1996 of embezzling between FFfr200 million and FFfr300 million from its funds (see *Nature* 379, 103; 1996 and *Nature* 382, 5; 1996).

The new research will be organized around five regional 'poles', concentrating on subjects ranging from gene function to tumour research. French public research agencies will manage the budgets of the different poles to ensure transparency.

► <http://www.arc.asso.fr>

Fish restaurant fingered by DNA data

Vancouver A database of salmon DNA has helped Canadian officials convict a Vancouver restaurant of serving up illegally caught fish.

Fisheries and Oceans Canada, the government body responsible for Canada's oceans and rivers, maintains a database of



Poached fish: chefs were undone by DNA.

25,000 samples of salmon DNA to enable researchers to monitor the spawning behaviour of the fish. The database's law-enforcement potential was revealed when officials used it to show that samples of salmon taken from La Baia Italian restaurant in the Vancouver suburb of White

Rock came from fish that had been caught in 1999 from the upper Fraser River near Vancouver. Fishing was banned in this part of the river at the time.

La Baia was fined Can\$8,000 (US\$5,200) after the owners pleaded guilty earlier this month.

What's in a name?

In the melting pot of modern science, chemistry's cutting edge is being rebranded as biology or nanotechnology. David Adam wonders if false modesty is leaving chemists to pick up the crumbs from their own periodic table.

Chemistry likes to style itself as the 'central science', but perhaps 'bridesmaid science' would be more appropriate. While other scientific disciplines reap maximum publicity from their triumphs, chemists have seen some of their brightest moments claimed by rival fields. From the discovery of lifesaving drugs to the explosion of work on carbon nanotubes, new developments in chemistry often seem to end up being appropriated by other disciplines.

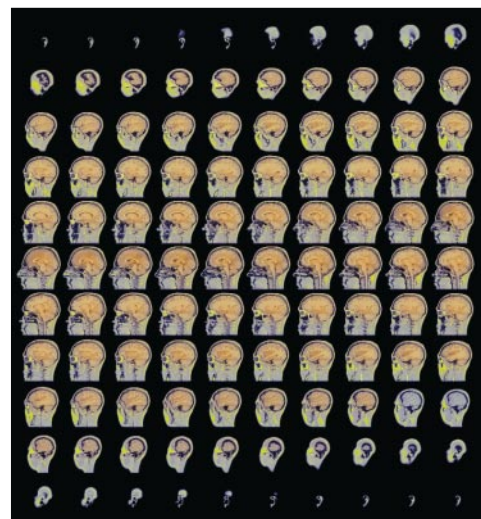
For researchers fighting to get the subject the recognition they say it deserves, chemistry simply does not smell as sweet by another name. And as chemistry's borders become blurred by collaborations with physicists and biologists, the researchers argue that it is time for the subject to claim achievements as its own — or risk failing to attract the brightest young scientists.

George Whitesides, a chemist at Harvard University, agrees that the most interesting parts of chemistry often end up being called something else. "It's something that I've never quite understood about chemists," he says. "They are in some profound way fairly modest. They develop these interesting techniques and then somebody else takes them and runs with them and nobody screams and hollers. You won't find a biologist who'll put up with someone taking what they're doing and walking off with it."

Unsung heroes

Whitesides' lab would be a good place to start the campaign to reverse this trend. He uses traditional chemistry techniques to investigate everything from how bacteria colonize surfaces to the design of fuel cells. One of the main current interests of his group is the synthesis of artificial cell membranes in the lab. Researchers can use these synthetic membranes to study how enzymes and sugars are adsorbed and recognized at the surface of living cells.

"If you think of what has happened in biology, a lot of it depends hugely on chemistry," says Whitesides. Early work on DNA, he points out, was done by chemists. "You wouldn't have DNA without the synthetic procedures for making nucleic acids and the gels to separate and run things out. Now that



Passing the buck: chemists discovered the existence of buckyballs (top right) as well as playing a vital role in the development of MRI scans (top left) — but other disciplines are now making the running.

ends up being called biology."

Chemists also played a vital role in developing nuclear magnetic resonance spectroscopy and its medical spin-off, magnetic resonance imaging (MRI). But today, MRI is often seen as an example of how physics can contribute to biomedical research.

Like the supporters of an impoverished lower-league soccer club, chemists have grown used to seeing their star performers transferred to more glamorous teams. Even Nobel prizewinning chemistry, it seems, does not stay chemistry for long. The discovery of the carbon spheres known as buckyballs was awarded the chemistry Nobel in 1996, but much of the work that has stemmed from that discovery is now seen as applied physics or nanotechnology.

But should chemists be worried by examples such as these? Collaborations are good for science, and chemists bring unique skills to these projects. Does it really matter if the

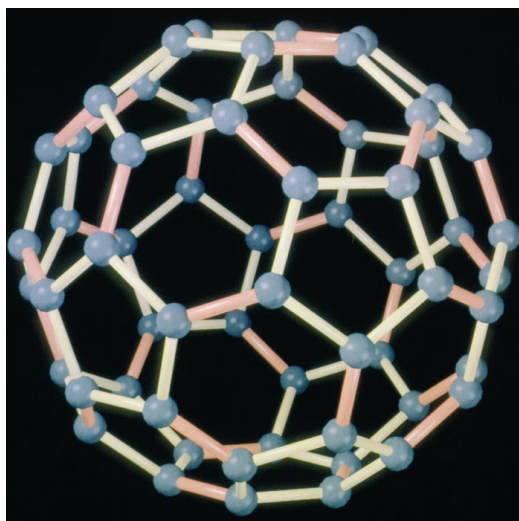
results get called something else? Harvard chemist Stuart Schreiber thinks it does.

"One could say it's a matter of perception or semantics," he says. "But perception and semantics can be important. You might be sending the wrong message to high-school students who are trying to determine which area of science they want to study in college."

Working closely with biologists in Harvard's Institute of Chemistry and Cell Biology, Schreiber's group synthesizes small organic molecules and uses them to study the functions of proteins. The molecules bind to proteins and alter their function — turning off some enzymes, for instance. Known as 'chemical genetics', the technique lets researchers examine the role of individual proteins without mutating the genes that code for them (see *Nature* **407**, 282–284; 2000).

"Now do we call that biology or chemistry?" asks Schreiber. "It's both. You're doing hard-core biology and you're doing hard-

TEK IMAGE/MAXIMILIAN STOCK LTD/MIT AT LAB/J. BERNHOLC/SPL



core chemistry. This approach wouldn't exist without the two integrated together."

So why is chemistry often seen as the silent partner in such collaborations? For Schreiber, it is the difficulty of explaining chemistry to non-experts. "One of the common features is that the research is initiated in a pure chemistry laboratory," says Schreiber. "The chemistry is complex, but the biology can often be described in terms a broader audience would understand."

As an example, he describes how chemists discovered a molecule that inhibits the enzyme HIV protease, which is critical to the AIDS virus's replication. New AIDS treatments followed. "*Time* magazine's man of the year was David Ho, a clinical researcher in AIDS," says Schreiber. "Did the public feel that this was a breakthrough in chemistry? Ho did a marvellous job, but the focus was so much on the clinical endpoint that it lost sight of where the science came from."

Schreiber and Whitesides say the drive towards interdisciplinary research is coming

from young researchers keen to stretch their wings. Some senior academic chemists are less enthusiastic, Schreiber says, which is one reason why the discipline has failed to brand its contribution to emerging areas.

Others argue that chemists tend to drop problems when they feel they are moving out of 'pure' chemistry. Michael Ward, a materials chemist at the University of Minnesota, says that chemistry research in universities has been restricted to certain areas. "Chemistry departments have isolated themselves by concentrating on what they perceive as pure science and discarding some research problems," he says. "Things like polymers and surfactants were picked up by chemical engineers and materials scientists, but that's changing now and chemistry is trying to get them back."

Heindirk tom Dieck, secretary-general of the German Chemical Society, adds that chemistry's contribution may be underestimated because it has not produced an eye-catching and expensive 'mega-project' like the human genome or space exploration.

Image problem

Whatever the cause of chemistry's image problem, the American Chemical Society (ACS) is seen by many as the body best equipped to remedy the situation. With more than 160,000 members, it is the world's biggest scientific organization.

"I think the ACS works tirelessly to create an appropriate and positive image of chemistry," says Schreiber. "On the other hand, I think many of the people involved have not been exposed to chemistry at the interface with modern neighbouring disciplines, and so may tend to understate or under-appreciate that component of our field. What's probably required there is a younger generation of spokespersons and writers who have

perhaps a better grasp of what is happening in the highly evolving aspects of chemistry."

The ACS says it is trying to catch up. A recent recruitment drive among younger chemists means that almost half of its members are under 50 for the first time in its 125-year history. "We've switched the demographic and the younger members are really helping to drive the society in the direction of being interdisciplinary," says Nancy Ryan Gray, director of the society's membership division. The ACS is also trying to promote chemistry's involvement in emerging disciplines by organizing conferences and publishing new journals on subjects such as proteomics and nanotechnology.

In Britain, the Royal Society of Chemistry (RSC) says it is broadening its membership criteria to allow for the recent growth in the use of chemistry research in biological areas. Researchers who do not have a classical chemistry background but work in 'chemical sciences' will be eligible to join, says Neville Reed, general manager for recruitment, members' services and communication.

Reed says this is part of a move to join up the dots to build up a stronger, more exciting picture of chemistry research. But he believes it is inevitable that some elements will be appropriated by other disciplines. "Chemistry has always been a service subject, it has always provided knowledge and skills that other people want to utilize," he says.

Others argue that not all collaborative projects are good for the subject. "Chemistry is trying to align itself with the excitement of genomics and all the promise that biology brings," says Julius Rebek, a chemist studying self-replicating molecules at The Scripps Research Institute in La Jolla, California. "But the chemistry involved is not leading-edge chemistry; it's pretty much worked out and understood."

Rebek believes the truly pioneering work on the boundary between the two disciplines is the attempt to create synthetic biological systems using chemistry. The big challenges, he says, are tasks like "making a living cell from scratch and making molecules that show primitive signs of life and recreating metabolism". For Rebek, this is pure chemistry. "After all," he says, "you wouldn't be a biologist until you succeeded."

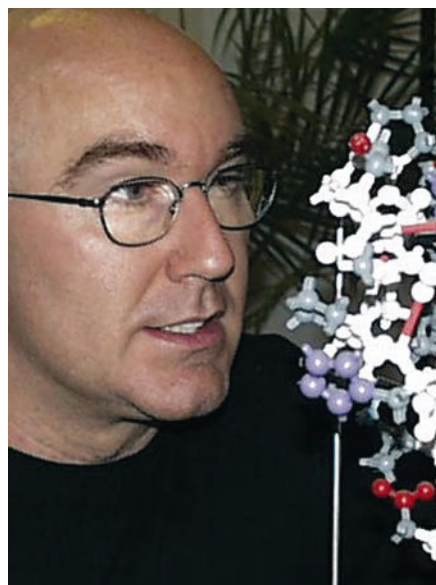
Whatever name is given to such projects, chemists will need to assert themselves if they want the results to be branded as chemistry. Scientific societies can play a role, but when it comes to the crunch, it is down to individual chemists to defend their territory. Perhaps it is time for them to lose a little of their modesty. "What we need is for chemists to talk up their own roles and be proud of what they do," says Reed. ■

David Adam is a news and features writer for *Nature*.

American Chemical Society ♦ <http://www.acs.org>

Royal Society of Chemistry ♦ <http://www.rsc.org>

German Chemical Society ♦ <http://www.gdch.de>



Claim reaction: Stuart Schreiber (left) and George Whitesides want to see chemistry get due credit.

Genetic medicine gets real

Gene therapists used to talk about permanently fixing 'broken' genes. But the emphasis has now shifted to treating conditions such as coronary disease and cancer using transient gene expression. Alison Abbott reports.

The language used by gene therapists has become noticeably more sober over the years. In the early 1990s, the field's pioneers raised hopes of curing single-gene disorders such as cystic fibrosis and severe combined immunodeficiency. Today, talk of cures has been replaced by guarded optimism about "encouraging" results. And the focus of attention has shifted away from genetic diseases towards using gene therapy as part of the therapeutic arsenal needed to confront major killers such as heart disease and cancer.

Two classical issues brought gene therapists down to Earth: efficacy and toxicity. The original idea of replacing missing or defective genes was compelling. But it soon became obvious that the vectors, such as weakened viruses, used to ferry the new genes into patients were highly inefficient. And in some instances, the vectors can also cause an adverse inflammatory reaction — as in the tragic case of the US teenager Jesse Gelsinger, whose death in September 1999 during a gene-therapy trial cast a pall over the entire field. Even when genes are successfully and safely transferred, their activity typically tails off after only a short period.

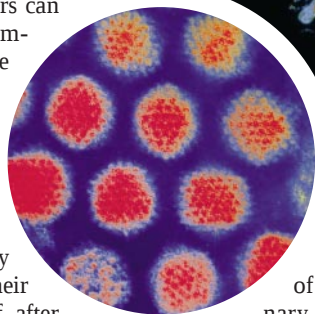
Faced with these problems, many gene therapists have gone back to the laboratory bench, trying to solve these technical issues in animal models before returning to human trials. But others have kept one foot in the clinic, turning to diseases that could be approached with the imperfect tools available. "We started to think positive," says Ronald Crystal of the Cornell Medical College in New York. "What applications could we have for transient gene expression?"

Growth potential

The list of contenders is growing, with cancer and coronary-artery disease at the top. Attempts to treat single-gene diseases now account for only about an eighth of the worldwide roster of 500 or so approved clinical protocols. Trials using transient gene expression — in which a short period of activity by transferred genes is sufficient



Gene ferries: adenoviruses (left), seen as red dots in the nucleus of this cultured human cell (above), can be used to introduce therapeutic genes into patients.



to treat a range of conditions — account for most of the rest.

Excitement over the potential of gene therapy to help treat coronary heart disease stems in part from work in Crystal's lab. In 1998, his team showed that introducing a gene encoding the protein vascular endothelial growth factor (VEGF) into the heart muscles of pigs caused new blood vessels to grow around blocked coronary arteries¹.

Crystal introduces his therapeutic genes using adenoviruses — a major cause of common colds — modified so that they cannot cause disease. But because adenoviruses are such common infectious agents, our immune systems are primed to recognize them, so the vectors are rapidly destroyed. This undermines attempts at efficient and permanent gene transfer which are needed to treat conditions such as cystic fibrosis — the focus of Crystal's earlier trials.

But the vectors' transience was no obstacle to Crystal's work with VEGF. "This is exactly the property we wanted for vascular disease," he says. "It's like flicking a master switch — once the blood vessels have been

triggered to grow, we don't need the switch to be flicked again." Indeed, excessive growth of blood vessels is something that Crystal wants to avoid.

Crystal has since moved into the clinic, testing the technique in 31 patients undergoing coronary-bypass operations in a phase I clinical trial, designed to assess safety². "Nothing about efficacy can be concluded with phase I trials," he stresses. Despite this, the patients reported some relief of symptoms, such as stress following exercise, and techniques for monitoring the growth of blood vessels indicated improvement in areas of heart muscle into which the vector had been injected.

Jeffery Isner of Tufts University in Boston is also working with VEGF. His phase I trial

The expression of transferred genes need only be enough to kick-start a response.

involved 85 patients with coronary heart disease and 110 with blockages in their peripheral circulation — a complication of conditions such as diabetes. Rather than using an adenoviral vector, Isner injects naked DNA encoding VEGF into muscle tissue in the vicinity of the blocked arteries. The efficiency of gene transfer in this method is very low, but Isner believes it will be sufficient to produce enough VEGF to trigger the growth of blood vessels.

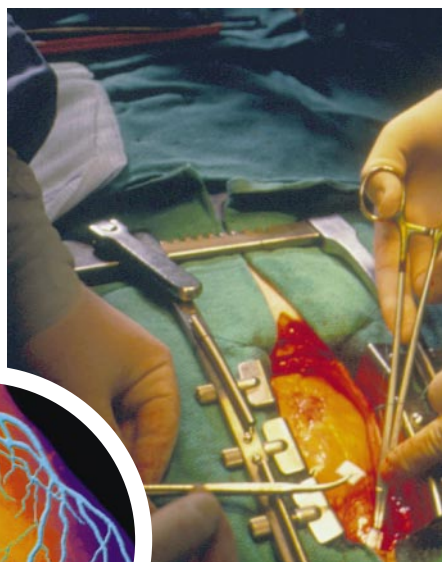
Preliminary results from Isner's trial, for 13 patients with coronary-artery disease, were published last August³. Again, scanning techniques suggested that the treatment had promoted the growth of new blood vessels.

In the wake of Gelsinger's death — in an unrelated gene-therapy trial at the University of Pennsylvania — progress has slowed. Both Crystal and Isner are now conducting further safety trials to be sure that the small number of deaths among very sick patients in their phase I studies occurred because of the patients' underlying health problems, rather than as a result of the therapy.

Stimulating work

Meanwhile, scientists working in the commercial sector are gaining ground. In March, at a meeting in Orlando, Florida, of the American College of Cardiology, Berlex Biosciences of Richmond, California, reported positive results with fibroblast growth factor (FGF), which also promotes blood vessel growth. In this trial, adenovirus vectors containing the gene for FGF were injected into the coronary arteries of 60 patients with angina. A further 19 patients served as placebo controls, receiving injections that did not deliver FGF. The treated patients enjoyed greater relief from their symptoms. At the same meeting, Berlex reported similarly encouraging results from a placebo-controlled trial in peripheral vascular disease. The company is now gearing up to start phase III studies, which will use large enough numbers of patients to determine efficacy.

Given these encouraging developments, gene therapists are now thinking about other conditions that might be treated using growth-factor genes. Gene therapy involving VEGF could, for instance, help wounds to heal by promoting the regrowth of skin and blood vessels. Several groups are looking at the potential of platelet-derived growth factor for the same application. Crystal is also conducting experiments on rats to test whether the genes for VEGF and BMP7, one of a family of proteins that promote bone growth, could be used to fuse vertebrae after spinal surgery. Currently, orthopaedic sur-



Replumbing the heart: Ronald Crystal has used transient gene therapy to stimulate the growth of new blood vessels around blocked arteries in pigs (top right, control and experimental). Now he is working with coronary-bypass patients.

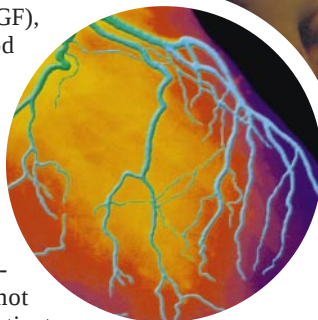
geons use metal rods and bone grafts. The gene for a signalling protein called Sonic hedgehog, which is involved in a wide range of development processes, may also have potential for transient gene therapy. One of the processes it helps to control is the development of hair follicles, and Crystal wonders if a short burst of Sonic hedgehog might be used to treat the hair loss caused by cancer chemotherapy, by activating dormant follicles. His team is investigating this by studying mouse pups, whose hair grows in synchronized bursts during their first weeks after birth. When Crystal and his colleagues injected adenoviral vectors containing the gene for Sonic hedgehog into the pups' skin, it shortened the 'resting' phases of the hair follicles⁴.

Tumour targets

While Crystal looks at treating a side-effect of cancer chemotherapy, other gene therapists are targeting the tumours themselves. Two-thirds of current gene-therapy

trials are for cancer, and around half of these involve immunotherapy, or 'cancer vaccines' — strategies that try to provoke the immune system's 'killer' T cells to attack tumour cells. Again, transient gene expression should, in theory, be sufficient to start the process. "The great thing about the cancer vaccine strategy," says Drew Pardoll of Johns Hopkins University in Baltimore, "is that expression of transferred genes need only be short-term, just enough to kick-start the immune response. Then the immune system should sustain it."

The immune system destroys many cells that go awry and start to turn cancerous. But tumours grow when it fails to notice that something has gone wrong. The main challenge facing the developers of cancer vaccines is breaking this immunological 'tolerance', says Pardoll. This requires the activation of specific T cells targeted to molecular signatures, or antigens, carried by the tumour cells.



T cells are activated by 'antigen-presenting cells', of which dendritic cells are the most important. Dendritic cells engulf and digest cellular debris, processing the resulting antigens so that they are carried on their surfaces. If the dendritic cells sense that these antigens represent a danger to the body, they also secrete 'co-stimulatory' molecules, and these activate the T cells that recognize the antigen involved. But triggering this sequence of events against antigens specific to tumour cells is a tough challenge.

One approach is to transfer a gene for a cytokine into tumour cells, and then to inject the resulting transgenic cells back into the patient. Cytokines are signalling proteins that help marshal immune responses, and previous research has shown that infusions of cytokines such as interleukin-2 (IL-2) can, by themselves, help shrink tumours. For example, in one trial of 255 patients with kidney cancer that had spread to secondary sites⁵, high-dose IL-2 caused remission in 12 patients, and a partial response in a further 24. But most patients experienced significant side-effects, particularly high fever.

Gene therapists are betting that using tumour cells to produce a transient burst of an appropriate cytokine will trigger dendritic cells more effectively, and so kick-start the chain of events to associate tumour antigens with danger — without causing generalized toxicity.

Early attempts at this strategy used tumour cells taken from the patients themselves. In 1998, for example, a team based at the Humboldt University in Berlin reported increased antitumour immune responses in 16 terminal melanoma patients vaccinated with their own cancerous cells transfected with genes for either interleukin-7 (ref. 6) or interleukin-12 (ref. 7).

But most experts believe that tailoring therapy to individuals by genetically manipulating the patient's own cells is too time-consuming to be a realistic clinical option. Instead, attention is now focusing on applying the same approach to cell lines maintained in culture that come from the same type of tumour as the patient's — which should share common antigens.

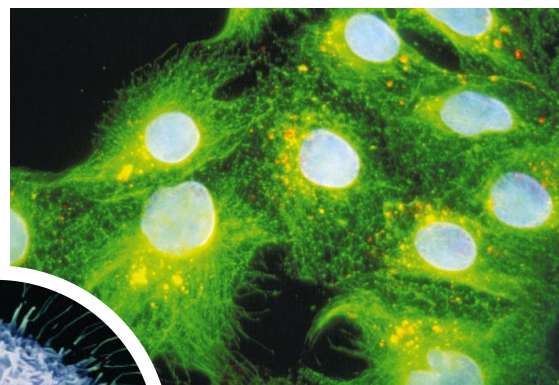
Early phase clinical trials are now being conducted by several companies, using the genes for cytokines including IL-2 and granulocyte-macrophage colony-stimulating factor. Prominent among these companies is Cell Genesys of Foster City, California, which is targeting melanoma plus kidney, prostate, pancreatic and lung cancers. Variations on the theme include transferring a second gene to the introduced cells to promote the desired immune response even further. These include the gene for a protein that binds to a molecule called CD40, which promotes the maturation of dendritic cells, or for co-stimulatory molecules.

In for the kill

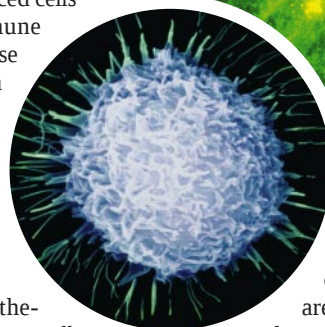
The disadvantage of 'off-the-shelf' vaccines using cancer cell lines is that some patients' tumours might carry a different set of antigens. In an attempt to avoid this problem, Vical, a company based in San Diego, is conducting phase I and II trials in which naked DNA encoding genes for cytokines or co-stimulatory proteins is injected into the patients' tumours. The idea is that the tumour cells will take up the DNA, and then transiently secrete proteins that will activate an immune response against the cells' antigens. The drawback, comments Pardoll, is that it is difficult to control the cells' uptake of DNA. Nonetheless, over the past two years Vical has reported at clinical research meetings encouraging results from patients suffering from melanoma, and prostate and kidney cancers.

An alternative approach to tackling tumours involves introducing the genes for cancer antigens inside viral vectors, in the hope that this will provide an appropriate context to associate the antigens with danger, and activate the immune system accordingly. Various early phase clinical trials are under way. Steven Rosenberg's team at the National Cancer Institute (NCI) in Bethesda, Maryland, for instance, has attempted to treat melanoma using adenoviruses carrying genes for specific antigens associated with the cancer⁸. Initial results were disappointing, however, with patients showing only weak immune responses.

In similar work, a team led by Jeffrey Schlom of the NCI and John Marshall from Georgetown University Medical Center in Washington DC, is tackling a range of cancers using a gene for a more generic antigen, called human carcinoembryonic antigen, spliced into the avipox virus^{9,10}. Avipox is a member of the vaccinia virus family that infects birds but which cannot replicate in mammalian cells. The immune response against the tumour antigen appears to increase with each monthly injection, particularly if the



Under attack: can prostate cancer cells be destroyed using 'cancer vaccines'?



injections also contain cytokines, and if the patients are first primed by vaccinating them with vaccinia.

At the 2001 meeting of the American Society of Clinical Oncology, held earlier this month in San Francisco, Schlom and Marshall's team announced the results of a small randomized trial of 18 patients with late-stage colorectal cancer, whose survival time was increased against controls.

At this stage, researchers in the field are unsure which cancer vaccine strategies will emerge as the most promising. "It's far too early to know which way is best," says Rosenberg, who reviewed progress in tumour immunotherapy in last week's *Nature*¹¹. "We need to try lots of different approaches just to see which works, and we need to work creatively," he says. Researchers in the field also stress that cancer vaccines are not likely to be used in isolation, but alongside surgery, radiotherapy and chemotherapy.

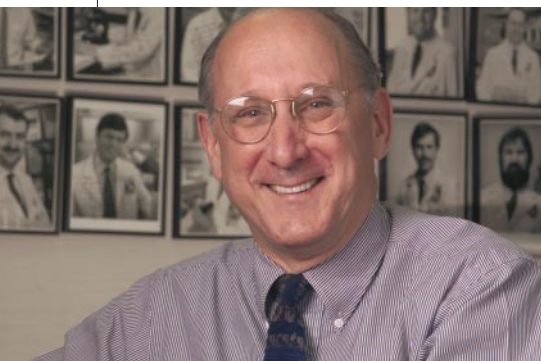
Indeed, although gene therapy approaches to both vascular diseases and tumour immunotherapy are showing promise, they are still in their infancy. The true efficacy of these treatments will only be revealed by definitive phase III trials, involving thousands of patients. But with companies such as Berlex Biosciences now planning such trials, we shouldn't have to wait too long to find out whether transient gene therapy will deliver the goods.

Alison Abbott is *Nature's* senior European correspondent.

1. Mack, C. A. et al. *J. Thorac. Cardiovasc. Surg.* **115**, 168–177 (1998).
2. Rosengart, T. K. et al. *Circulation* **100**, 468–474 (1999).
3. Vale, P. R. et al. *Circulation* **102**, 965–974 (2000).
4. Sato, N., Leopold, P. L. & Crystal, R. G. *J. Clin. Invest.* **104**, 855–864 (1999).
5. Fyfe, G. et al. *J. Clin. Oncol.* **13**, 688–696 (1995).
6. Moller, P. et al. *Br. J. Cancer* **77**, 1907–1916 (1998).
7. Sun, Y. et al. *Gene Ther.* **5**, 481–490 (1998).
8. Rosenberg, S. A. et al. *J. Natl Cancer Inst.* **90**, 1894–1900 (1998).
9. Marshall, J. L. et al. *J. Clin. Oncol.* **17**, 332–337 (1999).
10. Marshall, J. L. et al. *J. Clin. Oncol.* **18**, 3964–3973 (2000).
11. Rosenberg, S. A. *Nature* **411**, 380–384 (2001).

Information on gene-therapy trials

▶ <http://www.wiley.co.uk/wileychi/genmed>



Positive approach: Steven Rosenberg is using transient gene therapy to tackle melanoma.

Piecing together the history of our knowledge of chimpanzee tool use

Sir — Whiten and McGrew reported in *Correspondence*¹ a fascinating Liberian postage stamp from 1906, suggesting that its image might be the first documentation of digging for termites by a chimpanzee. They identified the stamp as printed in London, but did not know the basis on which the image was composed.

I have now solved the mystery. By chance I discovered the original drawing (see figure, right) that must have served as model for the Liberian stamp (figure, left), in a popular German book of the time². The artist is Gustav Mützel (1839–1893), whose brilliant illustrations of mammals and birds are well known through works such as those of the zoologist Alfred Edmund Brehm and his father, the ornithologist Christian Ludwig Brehm. Mützel signed the chimpanzee picture, adding the note “*n. d. Leben*” to show that his drawing was from his own observation of a living ape’s behaviour, not from a pelt, photograph or other illustration.

Whiten and McGrew are correct to praise the accuracy of the drawing, which is even visible on the stamp. But it is not of a wild ape. The figure shows the female chimpanzee Mafuka from Gabon, who lived in the early 1870s in Dresden zoo where Mützel drew her². Mafuka and other zoo apes learnt to use tools without any instruction: for example, drinking carefully out of a glass or a cup by imitating their human companions. These observations were among the reasons for establishing an ape research station on Tenerife, where Wolfgang Köhler conducted his famous studies on the intelligence of anthropoids³. In Mützel’s picture the ape is probably using a stick to explore a knot-hole in the trunk of a tree. But the natural-looking environment is added from the illustrator’s imagination. There is no termite mound in the picture, so Jane Goodall remains the first person to document wild chimpanzees ‘fishing’ with tools for termites⁴.

This outcome of the mystery of the stamp seems at first sight to be somewhat disappointing. Nevertheless, the story tells us something, not about chimpanzees, but about the breadth of human culture and global information flow. A nineteenth-century German artist portrayed a young ape in Germany and added attributes of the ape’s homeland to the scene. A British philatelic illustrator liked this African motif and transposed the scene from the German drawing to a Liberian stamp. Looking at the stamp nearly one



Tool order: did the stamp-illustrator know more about chimpanzees than the original artist?

hundred years later, two primatologists thought that it might be a record of tool use by chimpanzees in Liberia.

Ulrich Kattmann

Department of Biology, Geo- and Environmental Sciences, Carl von Ossietzky University Oldenburg, D-26111 Oldenburg, Germany

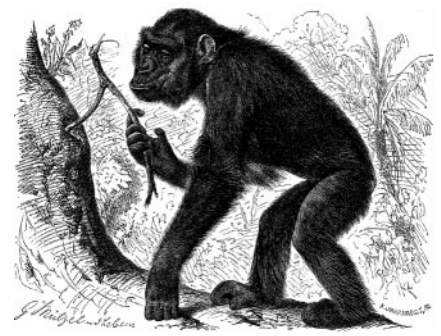
1. Whiten, A. & McGrew, W. C. *Nature* **409**, 12 (2001).
2. Brehm, A. E. *Brehms Thierleben: Erster Band* (2nd edn), (Bibliographisches Institut, Leipzig, 1876).
3. Köhler, W. *Intelligenzprüfungen an Anthropoiden I* (Königliche Akademie der Wissenschaften, Berlin, 1917).
4. Goodall, J. *Nature* **201**, 1264–1266 (1964).

Whiten and McGrew reply — We thank Kattmann for finding the elusive chimpanzee image used in designing the 1906 stamp that so intrigued us. This is a welcome demonstration of the power of international science, through the medium of *Nature*, to recover such an obscure item.

Kattmann suspects that the provenance of the 1887 image means we misinterpreted what the 1906 stamp itself portrays. Not necessarily. We recognized at the outset that the stamp design may have drawn on multiple sources of information, of which the image of the chimpanzee itself is only part. Indeed, closer inspection suggests that the 1906 stamp-designer knew more than the 1887 artist: the twig is shortened and no longer touches any object; the tree-hole and roots are replaced by what seems to be a mound; and the surface texture of the latter is different. The specificity of these changes suggests that the stamp design incorporates early knowledge of tool use in acquiring termites.

This would not have been the earliest report of tool use by chimpanzees, which we believe to be that concerning use of stone hammers to crack nuts in West Africa¹. Our suggestion was rather that here we might be seeing the first realistic depiction of tool use.

Of course, a single stamp illustration cannot prove this point. We believe, rather, that the close match with what we now know of termite-digging may not be coincidental. Although Kattmann’s rediscovery of the 1887 image certainly clears up some of the mystery, we remain keen to receive any information that



tracks down the stamp’s designer and the rationale for all he or she drew, for this might yet lead us to fuller and more informative early data on chimpanzee behaviour.

Andrew Whiten* and William C. McGrew†

**Scottish Primate Research Group, School of Psychology, University of St Andrews, St Andrews KY16 9JU, UK*

†Departments of Anthropology and Zoology, Miami University, Oxford, Ohio 45056, USA.

1. Savage, T. S. & Wyman, J. *Boston Journal of Natural History* **4**, 362–386 (1844).

Three-person rule was not set by Nobel himself

Sir — The Correspondences “Protest at Nobel omission of Moncada” (*Nature* **396**, 614; 1998) were a welcome contribution to the debate about criteria for selection of Nobel laureates. The editor’s note on the same page read: “The most troublesome aspect of the Nobel process is the apparently unchangeable fact that the prizes are distributed according to the terms of Nobel’s will, which states that the number of recipients in each category shall be limited to three.”

This clarification is not strictly correct. There is no reference to the ‘three-person rule’ in Nobel’s will. This principle was established in paragraphs 1b and 4 of the statutes of the Nobel Foundation that interpret and elucidate the will (see the Nobel Foundation website at www.nobel.se). “There shall be no departure from the following main principles ... In no case may a prize be divided between more than three persons.” The provisions in these paragraphs were the subject of lengthy negotiations from 1896 to 1900 between Nobel’s executors, representatives of his (largely disinherited) heirs, awardees of the prize and the Swedish government.

Lori Bradford, Perrie O’Tierney

Department of Anatomy and Cell Biology, Botterel Hall, Ninth Floor, Queen’s University, Kingston, Ontario K7L 3N6, Canada

Final frontier or ultimate mystery?

Plumbing the origins of human uniqueness.

A Mind So Rare: The Evolution of Human Consciousness

by Merlin Donald

W. W. Norton: 2001. 361 pp. \$28.95

John C. Marshall

A visit to my local bookshop the other day yielded 42 recent volumes with “consciousness” in their title (before I stopped counting). Merlin Donald, a professor of psychology at Queen’s University in Ontario, has managed to get both “consciousness” and a second buzz-word — “evolution” — into his subtitle. Have we really discovered so much about consciousness as to warrant the present deluge? The short answer is no. But it can, nonetheless, be entertaining to speculate on both the nature of consciousness and the reasons for our present obsession with the topic.

A Mind So Rare contains a very limited discussion of recent work on the cerebral correlates of visual awareness. There is some, not entirely reliable, discussion of attentional systems and the ‘binding problem’ (how the brain integrates information from distinct regions that each respond only to shape, colour, movement, depth and so on). The conclusions that can be drawn from functional brain imaging are also touched on, although one might, for example, hesitate before labelling certain aspects of the images produced with positron emission tomography (PET) as representing ‘positive flow’ and others as indicating ‘negative flow, or inhibition’. But Donald has other fish to fry. He has a well-justified distaste for radical reductionism, and expresses considerable scepticism about the credibility of the “neo-Darwinian Hardliners” who shout their “uncompromising belief in the irrelevance of the conscious mind and the illusory nature of free will”.

Although I share some of Donald’s prejudices in this regard, I was disappointed that he makes no attempt to address the claims made by the revival of sociobiology. And his attack on the ‘modularity of mind’ thesis seems insufficiently aware that there are good reasons why mental modules should operate in an encapsulated and automatic fashion: higher organisms simply do not have the time and resources to *think* about everything. In this regard, the brain operates on the same principle as much of our internal bodily workings — if something can be done on automatic pilot, let it be so done.

None of this is to deny that people do occasionally think and that consciousness might play a substantive, even central, role in what makes us truly human. The problem is



Flamenco! A creation of consciousness.

to work out how exactly this might be so. Donald’s claim is that “the human mind is unlike any other on this planet, not because of its biology, which is not qualitatively unique, but because of its ability to generate and assimilate culture”. Well, yes and no. It is hard to see how the human mind/brain could create cubism, quantum mechanics, King Lear, PET and rock-and-roll unless it were indeed qualitatively very different from the mind/brain of any other species on the planet. And the notion of culture has proved notoriously difficult to analyse in ways that go beyond common sense.

Donald has scant regard for the gene/meme analogy — memes being the propagation of ideas from one mind to another. According to this analogy, “memes are said to rise and fall, like nematodes and dinosaurs, according to the edicts of the Universal Acid of Darwin’s central idea, natural selection”. “Universal Acid” is the philosopher Dan Dennett’s expression for the quaint idea that natural selection (“survival of the fittest”) makes obsolete all traditional notions of truth and beauty. And what are

the grounds for Donald’s scepticism about memes? That the meme analogy makes the conscious mind “just a harmless voyeur, prone to delusions of grandeur about its ability to influence events”. But the concept of memes is surely just a bad joke, irrespective of any consideration of whether they ‘replicate’ with or without the intervention of consciousness.

Donald is no doubt correct to stress that there are severe limits to what the unconscious machinery of the mind can achieve. He reminds us time and again that “most learning requires the concentration of all our conscious resources on the task at hand”, and that the perception of our environment often demands that we direct “conscious attention toward the relevant aspects of the world”. For Donald’s purposes, it seems that consciousness is fundamentally much the same as attention. Yet the biological basis of attention is only marginally less mysterious than that of consciousness. We can observe that attention modulates the activity of brain cells, but exactly how this occurs remains unclear.

AP

What Donald does really well is to emphasize the huge size of the arena in time and space over which human attention can range. But it is perhaps significant that he has to illustrate the power of the conscious mind in culture by drawing on literature — Stendhal, Balzac and Henry James — and flamenco dancing, rather than science. Certainly, he makes little attempt to explain in detail how humans became so smart (and so stupid). Donald invokes such notions as “superplasticity” and “deep enculturation”, yet cannot make them sufficiently explicit to do the job. Similarly, the claim that no more than a larger brain distinguishes us from chimpanzees, and that brain size in itself allows for greater cultural development, does not seem overwhelmingly plausible.

A *Mind So Rare* is a good read, but leaves one little the wiser, in the main because Donald seems to be firing wildly in all directions. The English philosopher, writer and scientist George Henry Lewes claimed in 1878 that consciousness “designates an ultimate fact, which cannot therefore be made more intelligible than it is already”. He may have been right. ■

John C. Marshall is in the Neuropsychology Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Oxford OX2 6HE, UK.

More on the mind

To help guide the non-specialist through the thickets of the brain, the *Oxford Guide to the Mind*, edited by Geoffrey Underwood (Oxford University Press, 2001; £9.99, \$19.95, pbk), is a useful selection from OUP's *Companion to the Mind*, edited by Richard L. Gregory, with some new entries.

Good communications

Synapses

edited by W. Maxwell Cowan, Thomas C. Südhof & Charles F. Stevens
Johns Hopkins University Press: 2001. 792 pp.
\$69.95, £54

Michael D. Ehlers and Guoping Feng

In dealing with a discipline that ranges from molecules to mental states, books on neuroscience frequently lapse into the all-too-familiar categories of the broadly appealing but overly general and the highly accurate but immensely cumbersome. *Synapses* navigates between these twin poles with authority, providing a thorough yet thoroughly readable account of synaptic transmission.

Few structures in the nervous system have received as much attention as synapses, and for good reason, since it is at these specialized cell-cell junctions that neurons communicate with one another.

The process of nerve-cell communication, with its underlying structures and molecular mechanisms, is of fundamental importance to many biologists. Recent advances in this fast-moving area, as well as recognition by a prominent Swedish academy last year for neuroscientists Arvid Carlsson, Paul Greengard and Eric Kandel, make this book particularly timely and useful for all students of the brain.

Drawing on the expertise of recognized leaders in the field, the editors have compiled 15 chapters that skilfully relay the rich history and complexity of synapses. A fascinating introduction by Max Cowan and Eric Kandel chronicles the historical underpinnings of the synapse and sets the scholarly tone for the rest of the book. This chapter is a must for beginning students and seasoned scientists alike. From Galvani's description of “animal electricity” in the late eighteenth century, through the landmark studies of Cajal, Sherrington, Dale, Loewi, Eccles, Fatt and Katz, among others, the pantheon of synaptic transmission is brought vividly to life. This exceptional exposition sets the scene for the role of synaptic transmission in nervous-system function, so that the reader cannot wait to discover more about the complexities of the synapse. The subsequent chapters do not disappoint.

Synapses covers most of the core features of communication between neurons, from the mechanisms by which neurotransmitter molecules are released, through to synaptic plasticity, the process by which communication between neurons is either enhanced or weakened. This latter topic receives particular attention, and deservedly so, since plasticity of synaptic connections is an integral part of such complex functions as learning, memory and brain development.

In each chapter, concepts are organized into clearly defined sections, with descrip-

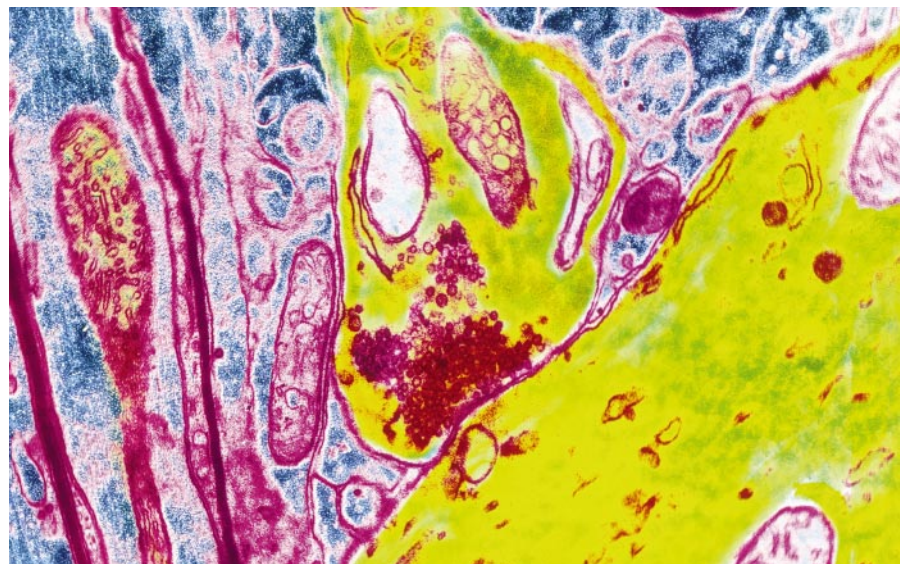
tions of cutting-edge topics presented as much as is possible in a book format. Literature citations for any given topic are thorough, and remaining controversies or ambiguities are discussed. The book's general organization is a march across the synapse: from the presynaptic terminal, across the synaptic cleft — the minute gap between the membranes of the two communicating neurons — and on to the postsynaptic membrane, with ample time for plasticity and development.

Despite the large number of authors, the tone is remarkably consistent and fluid. And considering the detail they contain, the chapters are surprisingly accessible. Those on postsynaptic specialization and synaptic plasticity are especially good presentations of current views, unifying a vast and often divergent array of experimental observations, and placing these concepts in the larger framework laid out by Cowan and Kandel.

Although the book primarily covers familiar topics, emerging fields are not ignored. A good example is the inclusion of a chapter on the synaptic cleft, an often neglected aspect. As Thomas Südhof points out, the synaptic cleft may hold the key to understanding how patterns of synaptic connections are formed during development, and therefore it deserves its chapter.

Almost certainly, *Synapses* will be used mostly as a reference, for it lacks the visual appeal of a course-oriented textbook, largely because of the long stretches of text-only pages. It will be particularly useful for neuroscientists who want to bring themselves up to speed (at least to *circa* 1999) on almost any aspect of synapse function.

As with any book covering a complex field, some topics get briefer mention than others, which to some extent reflects the interests of the authors. In particular, the



Information transfer: electron micrograph of a synapse, where one neuron meets another.

structure and function of neurotransmitter receptors could have been discussed in more detail, given their central role in synaptic transmission. Also, an overview of synaptic transmission in neurological and psychiatric disease would have broadened the book's appeal. *Synapses* nevertheless remains an impressive book that will be a useful reference to neuroscientists at all levels and a potential tool for teaching graduate students and advanced undergraduates. ■

Michael D. Ehlers and Guoping Feng are in the Department of Neurobiology, Duke University Medical Center, Box 3209, Durham, North Carolina 27710, USA.

Between

Liquid Crystals: Experimental Study of Physical Properties and Phase Transitions

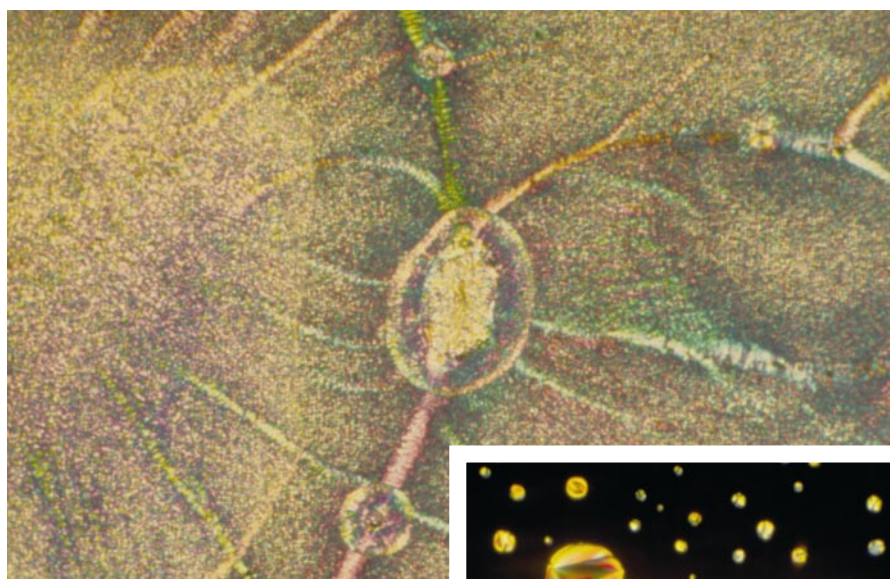
edited by Satyendra Kumar
Cambridge University Press: 2000. 483 pp.
£70, \$110

John Lydon

Any discussion of liquid-crystalline phases usually starts by describing them as a state of matter that is intermediate in structure between the more-or-less totally ordered crystalline solid phase and the completely disordered state of a liquid. This is a perfectly valid statement, but it carries all the wrong connotations: it implies that the properties of these 'mesophases' will also be intermediate. Admittedly, in some cases this is true. But in general, the combination of fluidity and spontaneous molecular ordering gives rise to unique and remarkable phenomena.

The liquid-crystal phases have, with good reason, been called nature's 'sensitive' phases. In fact, they are sensitive to virtually everything. Some change colour with temperature (and are sensitive to one-hundredth of a degree). Others respond to small electric fields and are readily aligned on treated surfaces, a combination of properties that makes them uniquely useful for display devices. Others form solvent phases that can align dissolved solute molecules, and this phenomenon is utilized in display devices and in new forms of spectroscopy. It is also possible to produce highly ordered, ultra-strong materials, notably Kevlar, by spinning fibres from liquid-crystalline polymer melts (rather than disordered liquids).

These unique characteristics are, of course, an expression of measurable physical properties, and over the past 20 years or so techniques for measuring them have featured increasingly in the research literature. The "long felt need" identified by Satyendra Kumar "for a source of general, non-super-

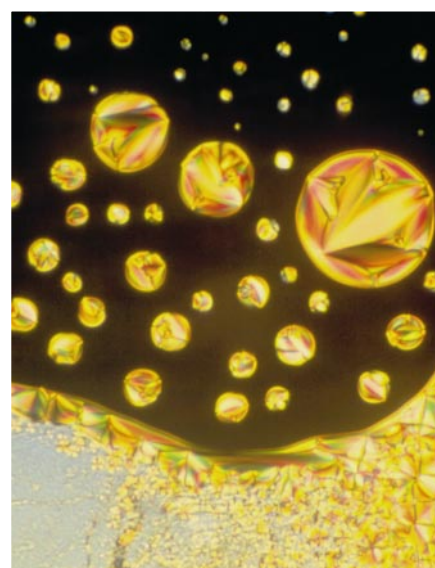


Fluid functions: liquid crystals examined by light (right) and polarizing (top) microscopy.

fluous and practical information regarding the experimental research techniques and how they are applied to liquid crystals" is real enough, and the book he has put together is most welcome.

Liquid Crystals competently and elegantly covers the principal techniques used to investigate liquid crystals — these include polarizing microscopy, differential scanning calorimetry, X-ray diffraction, nuclear magnetic resonance and light scattering. And the chapter on physical properties covers all the familiar themes — dielectric properties, optical diamagnetic elastic and viscous properties, electrical and thermal conductance and density measurement — with 20 pages of colour plates showing the main optical textures. There is a remarkable chapter on freely suspended films — films stretched across a hole in a plate so that they are not in contact with the solid surface on either side. And the book ends with a discussion of the relationship between chemical structure and mesogenic properties. Each topic receives a mathematical treatment, at a level comprehensible to a physics graduate.

The patchwork of chapters by a dozen authors is almost inevitably selective and idiosyncratic, and in many places the book has depth but not breadth. Esoteric points are sometimes discussed in detail. For example, the authors from Ohio's Kent State University are proud of their characterization of the elusive biaxial nematic phase and could not resist mentioning it in the introduction. But there is virtually nothing about lyotropic phases, polymer liquid crystals or discotics. Nevertheless, Kumar has achieved his goal of producing a volume of practical information with "in-depth technical details". For students about to embark on experimental investigations of thermo-



tropic, nematic or smectic systems, this book could be one of their best buys. ■

John Lydon is in the Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9TJ, UK.

Flitting around the subject

Fly: An Experimental Life

by Martin Brookes

Weidenfeld & Nicolson: 2001. 215 pp. £16.99

Peter Lawrence

It is a 'truth' universally acknowledged that in any popular piece of science writing or television, the science must be diluted with entertaining asides. Otherwise, it is feared, the reader will doze off in her chair and the viewer change channel. Martin Brookes appears to have embraced this belief when he set out to write about the fruitfly *Drosophila*, about the history of research using it, and about the fly's place in our present understanding of biology. Naturally, he has to communicate some facts and ideas. He fears, however, that his readers will find science a bore, so he offers us



EYE OF SCIENCE/SPL

Fruits of experiment: a mutant fly with two small extra eyes in place of antennae.

little holidays from the hard work as we go along. The result? A book written as if by a distract and garrulous aunt who cannot stick to the point for more than a few seconds — it fizzles with the insubstantial.

Brookes sets the scene with the prose of a fancy menu. Do flies feed? No — “their digestive juices work chemical magic on a belly full of banana goo”. Imagination unleashed, he then gives us a sketch of T. H. Morgan, the greatest of all developmental geneticists, whose wiry beard contained “a morsel of rotting banana ensnared in the hooks of his facial hair”. Then something a bit more substantial — a discursive mini-biography of Morgan and a picture of his world, apparently using Garland Allen’s fine biography, *Thomas Hunt Morgan: The Man and his Science* (now, sadly, out of print), as the source; this is informative and interesting. There are titbits about Morgan’s life with sea spiders and worms, and how he met the fruitfly — as well as his wife.

Brookes has a lively tabloid style and can sometimes be sharp and even unfair — as in his summary of Lamarck’s theory of evolution: “It was all very complicated. It was also complete garbage.” He does better when he sticks to a theme. Morgan’s discovery of the *white* mutation and how he deduced its location on the X chromosome is well described.

The heart of this brief book is a set of unrelated chapters that sample some of the topics that *Drosophila* has illuminated. “Unscrambling the egg” is about the fly’s huge contribution to our knowledge of development. Here Brookes continues to entertain, but his writing betrays a lack of understanding of the subject, which gives the whole a very second-hand feel.

He then turns his attention to evolutionary genetics and, in particular, to the life and contributions of Theodosius Dobzhansky. He clearly knows much more about this subject and could have conveyed more understanding. But he cannot stick to a storyline, and all of us — author, reader, Dobzhansky

and Morgan — are picked up and spun around like dust in a desert whirligig. Next comes behavioural genetics, where Brookes appears to have depended largely on Jonathan Weiner’s lyrical book *Time, Love, Memory* (Knopf, 1999). And just in case he might be accused of getting heavy, we have short and spicy chapters on sex, ageing and speciation.

But the important question is: who is this book for? Not for the likes of me, perhaps, for I have grown accustomed to the little faces of fruitflies. So I gave it to a first-year biology undergraduate. She read the book quickly and easily, and liked its lively narrative style — a definite plus there. But she also wondered who would want to read it. Not her, she felt: when the material covered was central and important, it did not inform but rather perplexed. For example, having been taught the basis of Mendel’s ratios in school, she found Brookes’s analogy of mating houses with black and white doors confusing. And when the information was more peripheral, such as a summary of Ulrike Heberlein’s approach to alcohol addiction using flies, she thought that the attempt to be hip and happening meant that the message itself got lost.

So my friend was amused rather than informed; perhaps this is not too bad, for she might now go and look for more information in heavier texts. Nevertheless, a scientific story can be exciting; history and logic alone can engage a reader, and I think Brookes would have done better had he explained more and embellished less. All the work on *Drosophila* over the past 100 years has provided much to help us understand ourselves and nature; but we still await a good popular book that explains just how. ■

Peter Lawrence is at the MRC Laboratory of Molecular Biology, University Medical School, Hills Road, Cambridge CB2 2QH, UK.

Chemical reflections

Miroir de la Chimie

by Pierre Laszlo

Seuil: 2000. 328 pp. 21.04 euros

Jacques Reisse

The organic chemist Pierre Laszlo writes extremely well, with a rare elegance. *Miroir de la Chimie* is a collection of short essays, in French, on a wide range of topics related to chemistry that will interest both the professional chemist and the non-chemist. The patchwork format and the arbitrary grouping of the essays into ‘themed’ chapters are a little off-putting at first, but if you treat the book as something to dip into, reading it becomes a real pleasure. Laszlo presents a succession of glimpses into

various aspects of chemistry and into the lives and activities of chemists, from obscure postdocs to Nobel prizewinners.

You might start with one of the essays in the first chapter — “Six faces of the molecular object” — but could equally well begin with “Six characters: a portrait gallery”, “Six themes: the chemical theatre”, or “Six instants in the research career”. The concluding “Bibliographic Essay” provides much useful information not only for the non-specialist but also for teachers of chemistry.

One of the book’s great strengths is the fact that the information provided is accurate and well documented. Laszlo is a careful observer of the scientific community to which he belongs and describes it perceptively but without malice. During his career, he has worked with or met some of the most prominent chemists in structural chemistry and stereochemistry. And over the past 40 years he has accumulated a rich hoard of observations not only about chemistry and its developments but also about chemists themselves and the social group that they make up. Like all human societies, this one has its characteristics, rules, intellectual leaders and many obscure workers, and Laszlo makes it come alive for us in all its colourful variety.

Laszlo knows that chemistry has a poor public image but he also knows how important it is to the world. He makes an impressive attempt to convince the reader of this, and deserves to succeed. One of his strategies here is a short but convincing piece of science fiction that describes what might happen if a large meteorite hit the Earth near the Gulf of Mexico, the site of many US petrochemical plants.

The book is really very good. It gives the non-chemist a realistic view of the chemical sciences and describes how the community works, how chemists interact with one another, how some of them achieve success, and the many chemists who have contributed, at different levels, to the dramatic development of the field. Laszlo’s book will also be appreciated by the chemistry community, particularly those interested in the history of their discipline, the role of leaders in the field, and the impact of major books, prestigious conferences and scientific prizes on the evolution of chemistry during the second half of the last century. ■

Jacques Reisse is at *Chimie Organique* (CP 165/64), Université Libre de Bruxelles, 50 Avenue F. D. Roosevelt, B-1050 Bruxelles, Belgium.

Correction

The photograph on page 519 of the 29 March issue (from the book *Glorious Eclipses*) depicts the Blue Mosque (Sultan Ahmet Mosque) in Istanbul and not Hagia Sofia as noted in the book and in our caption.

Explanatory tyranny

If an explanation seems wonderfully simple, it's probably too good to be true.

Timothy Taylor

As an archaeologist, I was not surprised that the human genome comprised far fewer genes than predicted. It had been assumed by some that genes for simple things, such as hair colour, would soon be joined by genes for other things, such as (infamously) homosexuality, also construed as 'simple' by those with inadequate knowledge of the natural history of sexual culture. Belief in a complex causal nexus, not just between genes but between genotype and environment — including the cultural environment, whose profound modifiers are known to extend even to aspects of the intra-uterine environment — seemed somehow less scientific than a search for discrete, function-specific units. Perhaps I have no business commenting on genetics, but what I am really interested in is an explanatory imperialism which threatens to subsume even archaeology. The unmet expectation in the genome case signals a wider epistemological malaise affecting public perceptions of what science is and how it should look.

Twenty-five years ago, Richard Dawkins, in *The Selfish Gene* (Oxford Univ. Press, 1976), proposed the existence of cultural replicators or 'memes' that worked in a similar way to genes. According to this idea, material things (such as mugs), aesthetic objects (such as Beethoven's *Ninth Symphony*) and ideas (such as God) were copyable codes; "the computers in which memes live are human brains". Dawkins considers his unit good, as it subsequently successfully replicated itself as a word. Indeed, in respect of brief snatches of music or simple and peculiarly uniform artefacts, such as coins, the idea seems to have some applicability.

But in aesthetic philosophy, musical performances are considered tokens of types: each performance of Beethoven's *Ninth Symphony* is unique and can never be reproduced. And in archaeology, artefacts are not generally considered to constitute species, successful or otherwise. Unlike genes, they

cannot be classified so that any particular 'defining' attribute is both sufficient and necessary for group membership (monothetic classification). In modern Britain the name 'mug' is given to a variety of forms, and no single attribute is both sufficient and necessary for group inclusion (polythetic classification). Whether a mug is able to hold boiling liquid (or not: a beer mug), or made of pottery (or not: a tin mug), or whether a thing with a handle is a mug (rather than a cup or a door), fuzzily constrains how any one object in our culture gets its designation and defined function, connotations, and symbolism. Moving beyond this familiar kind of complexity, anthropologists must attempt to tackle alien classifications, while archaeologists face a further level of difficulty in confronting alien and extinct cultural phenomena. Prehistoric archaeology attempts to observe and explain these without the aid of texts and is therefore — for all its 'the spade never lies' appeal — an exacting philosophical challenge (spades never speak either).

Mugs cannot replicate themselves and memes have not been embraced by cultural theorists. So my protest would be unnecessary were it not for the baleful influence of 'gene-meme' thinking in the recent spate of para-archaeological fantasies which purport to reduce the very great complexities of the human past to a simple formula. Theories that attribute pyramids in both Egypt and Mesoamerica to a uniquely gifted race of mystic astronomers — whether priests of Ra, ancient Antarticans, or aliens — migrating across continents bringing true culture, have been curiously dignified by the meme idea. Likewise, claims for a standard, memic, unit of measurement, the 'megalithic yard', in neolithic Orkney and Egypt — the secret intellectual property of an international masonic guild — find converts by fulfilling the public's formal expectations of science.

Many architectural structures are broader at the bottom than at the top rather than vice versa, and for good reason. Call them pyramids, but the Giza mausoleums, originally smooth and unclimbable, differ fundamentally from the Mayan temples of 3,000 years later, where ceremonial staircases led to elevated sacrificial platforms. Messy but powerful bayesian methods reveal the chimaera of the 'megalithic yard' too. Layout measurements were based on the human body — hands, feet, paces — which can explain the metrological convergence and coincidence across place and time.



All pyramids? The Mayan temple above differs fundamentally from those at Giza.

Ever since Plato there has been the idea that good explanations should condense into simple formulae. Umberto Eco, in *Foucault's Pendulum* (Secker & Warburg, 1989), says "The truth must be simple and sayable: $E = mc^2$ "; knowing full well that things are seldom like that in culture. Oscar Wilde is not just witty, but humane, to observe that "the truth is rarely pure, and never simple". The proper explanation of archaeological phenomena cannot look like $E = mc^2$. Archaeology has to explain its subject in terms of 'how ... possibly [did that event occur in the past]' (h-p explanation) rather than 'why ... necessarily [does that always happen in the lab]' (w-n explanation). This reliance on h-p explanation is shared with other systematic endeavours to explain ontologically unique, unrepeatable phenomena, such as the conditions for the creation of the observable Universe or the extinction of the dinosaurs (whether or not these appeal to certain covering laws).

The current ascendancy of the w-n form is a response to the perception that the public and politicians need clear leads on how research is conducted. But it is also supported within science, by individuals and committees. Yet all scientific enquiry possesses some historical, ontologically unique dimension, and what counts as secure knowledge in the explicitly historical sciences is complex, multi-stranded and, within limits, indefinite. Attempting to shoe-horn explanation into a one-size-fits-all form represents nothing less than the dumbing-down of explanation. ■

Timothy Taylor is in the Department of Archaeological Sciences, University of Bradford, Bradford BD7 1DP, UK.

What counts as secure knowledge is complex, multi-stranded and indefinite.

A question of scale

Tamas Vicsek

If you search the Internet for the phrase 'collective behaviour', most of what pops up will be about group activities of humans, including riots, fashion and mass panic. But collective behaviour is also an important aspect of observed phenomena in the world of atoms and molecules, for example during spontaneous magnetization. In your web search, you might also find articles on collectively migrating bacteria, insects or birds; or phenomena where groups of organisms or non-living objects synchronize their signals or motion — think of fireflies flashing in unison or people clapping in phase during rhythmic applause.

Is there a common factor in these seemingly diverse phenomena? The answer is yes — they happen in systems consisting of many similar units interacting in a relatively well-defined manner. These interactions can be simple (attraction/repulsion) or more complex (combinations of simple interactions) and can occur between neighbours in space or in an underlying network. Under some conditions, transitions can occur during which the objects adopt a pattern of behaviour almost completely determined by the collective effects of the other objects in the system.

The main features of collective behaviour are that an individual unit's action is dominated by the influence of its

neighbours — the unit behaves differently from the way it would behave on its own; and that such systems show interesting ordering phenomena as the units simultaneously change their behaviour to a common pattern.

Ferromagnets are a good example. These materials can undergo spontaneous magnetization, in effect because they are made up of a host of 'tiny magnets'. At relatively high temperatures, these small magnets cannot align with each other and the resulting magnetization is zero. But at a critical temperature, the tendency to adopt a common direction suddenly but continuously takes over from the effects of fluctuations. So most of the small magnets, assisting each other in a collective manner, point in the same direction, and magnetization spontaneously appears. Similarly, a group of feeding pigeons randomly oriented on the ground will order themselves into a uniform flock as they fly away after a big disturbance.

Collective behaviour applies to a great many processes in nature, which makes it an extremely useful concept in many contexts. Understanding a new phenomenon is usually achieved by relating it to a known one: a more complex system is understood by analysing its simpler variants. In the 1970s, there was a breakthrough in statistical physics in the form of the 'renormalization group method', which gave physicists a deep theoretical understanding of a general type of phase transition. The theory showed that the main features of transitions are insensitive to the details of the interactions between the objects in a system. This means that, as an extreme case, the orientational forces between atoms can result in ordering phenomena similar to those seen in groups of much more complex units.

Consider, as a thought experiment, thousands of people standing on a square and trying to turn to face the same, but externally unspecified, direction after being asked to do so. A nice example of collective behaviour would be if all of them managed to face the same way. But can they do it? Statistical physicists can predict for certain that they cannot. The physicists base this on a theorem valid for particles with short-range ferromagnetic interactions, which states that in two dimensions no long-range ordered phase can exist for any finite temperature and zero external field. So what happens in our example? On a small scale, people are looking almost in the same direction, but on a large scale, as seen from a helicopter for example, they form vortex-like directional

Collective behaviour

The way in which an individual unit's activity is dominated by its neighbours so that all units simultaneously alter their behaviour to a common pattern.

patterns because of the small perturbations caused by human errors — just like the little magnets. But if people in the crowd are allowed to choose from a few discrete directions, the required ordering can be achieved.

Another recent theory surprisingly predicts that if the people are asked to move in the same spontaneously selected direction they can do it. Simulations and analysis of the related equations show that the motion of the individuals results in additional interactions, which enables information about local directions to spread through the whole system.

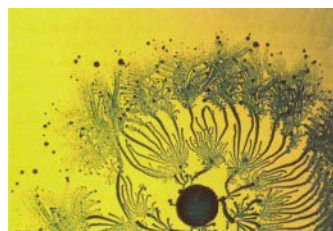
The approach of treating flocks, or even crowds, as systems of particles naturally leads to the idea of applying successful methods of statistical physics, such as computer simulations or theories on scaling, to the detailed description of the collective behaviour of organisms. Indeed, for the past couple of years — making use of collective behaviour — many interesting discoveries have been made for a rich variety of physical, animal and human behaviours.

We now have a much better understanding of collective phenomena such as aggregation, swarming and synchronization. For example, modelling the Internet's evolution shows that it is less sensitive to random, but more vulnerable to intentional, attacks than many other kinds of networks. Quantitative prediction of the collective reaction of people to situations including panic, and vehicular or pedestrian traffic is already possible in some cases. Using computer models to simulate cars or people as particles with appropriate interactions is soon expected to lead to better design of highways or hypermarkets. An improved, collective behaviour-based numerical analysis of situations involving higher-level human activities, such as fashion or outcomes of elections, may become feasible in the not-too-distant future. ■

Tamas Vicsek is in the Department of Biological Physics, Eötvös Loránd University, H-1117 Budapest, Pazmany P. Stny 1A, Hungary.

FURTHER READING

◆ <http://angel.elte.hu/~vicsek>
Coleman, J. S. *Foundations of Social Theory* (Belcamp Press, Cambridge, MA, 1990).
Huang, K. *Statistical Mechanics* (Wiley, New York, 1987).
Vicsek, T. (ed.) *Scaling and Fluctuations in Biology* (Oxford University Press, Oxford, 2001).



All together now: collective behaviour can readily be seen in fish and bacteria (inset).

Almost Planet X

S. C. Tegler and W. Romanishin

Optical and infrared observations of a bright object in the outer Solar System reveal it to be surprisingly large — almost as big as Pluto's moon. It could be the first of many such discoveries.

When Neptune, the eighth planet in the Solar System, was discovered in 1846, astronomers began looking for a ninth. Inspired by irregularities in Neptune's orbit, Percival Lowell began searching in 1905 for what he called Planet X. But it wasn't until 1930, 14 years after Lowell's death, that Clyde Tombaugh of the Lowell Observatory in Flagstaff, Arizona, discovered Pluto. The search for Planet X was over. Or was it? Pluto turned out to be much smaller than expected, and some astronomers, including Tombaugh, continued to search for a tenth planet, beyond Pluto.

No one has found a tenth planet, but during the past decade astronomers have discovered that Pluto is not alone. It seems that Pluto and its moon Charon are the largest known members of an ancient ring or belt of icy bodies, known as the Kuiper belt^{1,2}. The primitive nature of the material in this belt might hold important clues about the formation and evolution of the Solar System. Unfortunately, apart from Pluto and Charon, these objects are so faint that studying their most basic properties requires a herculean effort. On page 446 of this issue, Jewitt, Aussen and Evans³ present the best measurements yet of the size and reflectivity of another body in the belt. Their work raises the possibility that Pluto is not the only Planet X, but perhaps one of several.

The body studied by Jewitt and colleagues, now called Varuna⁴ (Fig. 1), was detected last November by the Spacewatch telescope in Arizona. It is the third brightest object in the belt, after Pluto and Charon. The optical and infrared measurements made by Jewitt *et al.* show that Varuna is slightly smaller than Charon, making it the third largest known object in the belt, but with a much darker surface than Charon. The results suggest that Pluto and Charon are not uniquely large objects, and that a continuum of sizes may exist. We can now imagine that bodies even larger and more distant than Pluto will be found. Such objects have so far escaped detection because of their extreme faintness, which is due in part to the feeble illumination from the Sun, and in part to their very dark surfaces.

The idea of a belt of icy bodies surrounding the outer planets goes back to 1930, soon after the discovery of Pluto. Leonard⁵ was the first to publish the idea, which was later expanded upon by Kuiper⁶ and Edgeworth⁷.

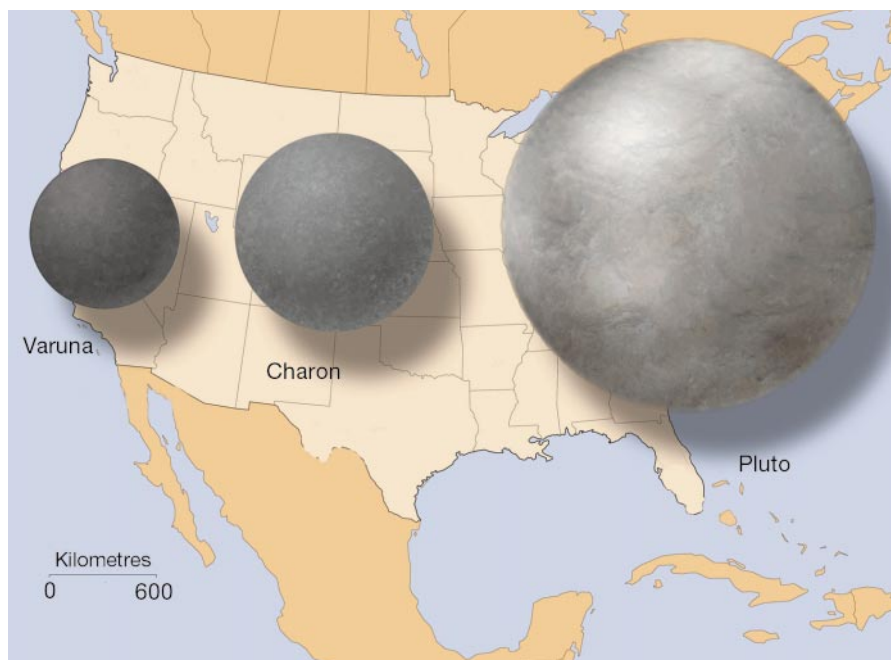


Figure 1 The size and surface darkness of the ninth planet Pluto, its moon Charon, and the newly discovered object Varuna, the three largest known members of an ancient belt of icy objects that surround the eight planets closest to the Sun. Varuna was discovered by the Spacewatch team at the University of Arizona⁴, who regularly scan the skies for asteroids and comets. Infrared and optical measurements by Jewitt and colleagues³ suggest that Varuna is by far the largest body so far discovered in the belt, apart from Pluto and Charon. This work implies that Pluto and Charon are not unusually large, and that there may be other, as yet undiscovered objects in the outer reaches of the Solar System, perhaps including 'planets' even larger than Pluto.

For 60 years, Pluto and its moon Charon were the only known objects in the Kuiper belt. The third was found in 1992 (ref. 1), following advances in telescope and detector technology. There are now nearly 400 known objects in the belt, although there may be hundreds of thousands of 100-km-sized bodies and possibly billions of 10-km-sized objects that have yet to be discovered⁸. Although there are so many Kuiper belt objects (KBOs), the total mass of the belt is probably only about a tenth the mass of the Earth⁸. All the evidence to date suggested that Pluto was a uniquely large object within the Kuiper belt.

Presently, the best way to estimate the diameter of a KBO is to measure the amount of sunlight it reflects, using an optical telescope. At a given distance, the brighter the object, the larger the surface area facing the Earth. But there is a caveat. The brightness of an object depends not only on the area of the reflecting surface, but also on the albedo,

the percentage of sunlight the surface reflects. Material as black as charcoal has an albedo of about 4%, whereas material as reflective as frost has an albedo of about 40%. So optical brightness measurements alone cannot distinguish between a large, dark object and a small, highly reflective one. Determining independent values for the diameter and albedo requires measuring both the sunlight reflected and the thermal infrared light emitted by the body.

This technique has been successfully used to determine the diameter and albedo of asteroids in the inner Solar System. But KBOs are much harder to detect in the infrared because they emit radiation at wavelengths that can be easily absorbed by the Earth's atmosphere. Some objects, known as Centaurs, are thought to have escaped from the Kuiper belt and can be found throughout the outer Solar System, sometimes passing even closer to the Sun, not far from Earth⁸. As the Sun warms them, some of the icy

material evaporates in a cloud of dust and gas, which is occasionally visible as the tail of a comet. Until now, astronomers have assumed that KBOs have an albedo similar to the closer Centaur objects and comets⁸, whose albedo is typically 4%. Such an assumption is fraught with danger because the Sun's warmth may trigger chemical and physical changes in Centaurs that alter their surface reflectivity.

Varuna is so bright that Jewitt and colleagues have been able to measure simultaneously its optical reflectivity and infrared emission; they found that it has a diameter of 900 km and an albedo of 7%. Its surface is darker than that of Pluto and Charon, but its albedo is slightly higher than the value previously assumed for KBOs. Only Pluto and its satellite Charon have larger diameters than Varuna (2,400 km and 1,200 km, respectively). So Varuna closes the gap between the largest previously known KBO, which was around 600 km in diameter (assuming an albedo equal to that of Varuna), and the smallest planet in the Solar System, Pluto.

Varuna and the handful of Centaurs whose albedos have been measured all have dark surfaces, which tells us there is little surface frost or ice. We already know that Pluto and Charon have very bright surfaces⁹. Because of its size, Pluto is a very different beast to the other KBOs — it is large enough to retain a tenuous atmosphere, resulting in a global surface frost. Charon, on the other hand, has a similar diameter to Varuna, but its albedo (40%) is about six times that of Varuna. Why do Varuna and Charon have

such different albedos? It could be because the surface of Charon is covered in frozen water¹⁰, perhaps a result of whatever process caused it to become a moon of Pluto.

More measurements of thermal emission are needed to determine whether dark surfaces are ubiquitous in the Kuiper belt, as well as to obtain reliable diameters. Thermal infrared observations of these faint objects are extraordinarily difficult from the ground, but much easier in space. The Space Infrared Telescope Facility (SIRTF) satellite is expected to measure the diameters and albedos of dozens of KBOs after its launch in 2002. Such observations are our best chance of finding more planets beyond Pluto. ■

S. C. Tegler is in the Department of Physics and Astronomy, Northern Arizona University, Flagstaff, Arizona 86011, USA.

e-mail: stephen.tegler@nau.edu

W. Romanishin is in the Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma 73019, USA.

e-mail: wjr@phyast.nhn.ou.edu

- Jewitt, D. & Luu, J. *Nature* **362**, 730–732 (1993).
- <http://cfa-www.harvard.edu/cfa/ps/lists/TNOs.html>
- Jewitt, D., Auel, H. & Evans, A. *Nature* **411**, 446–447 (2001).
- McMillan, R. & Larsen, J. *Minor Planet Electronic Circular 2000–X02* (Minor Planet Center, Cambridge, MA, 2000).
- Leonard, F. C. *Leaflet Astron. Soc. Pacific* **1**, 121–124 (1930).
- Kuiper, G. P. in *Astrophysics: A Topical Symposium* (ed. Hynek, J.) 357–427 (McGraw-Hill, New York, 1951).
- Edgeworth, K. E. *J. Br. Astron. Assoc.* **53**, 181–188 (1943).
- Jewitt, D. C. & Luu, J. X. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 1201–1229 (Arizona Univ. Press, Tucson, 2000).
- Tholen, D. J. & Buie, M. W. in *Pluto and Charon* (eds Stern, S. A. & Tholen, D. J.) 193–219 (Arizona Univ. Press, Tucson, 1997).
- Marcialis, R. L., Rieke, G. H. & Lebofsky, L. A. *Science* **237**, 1349–1351 (1987).

(the nucleotide-binding motor domain) of kinesin^{2,3}. So it had been suspected that these two families of motor proteins operate according to similar principles⁴. This idea is reinforced by the findings of Kikkawa *et al.*¹.

In both myosin and kinesin, the nucleotide-binding site is flanked by two structural elements, referred to as switch I and switch II, by analogy to equivalent loops found in G proteins (signalling proteins that bind to GTP). The crystal structures of myosin and G proteins in different nucleotide-binding states (that is, bound to GTP or GDP for the G proteins, and to ATP or ADP for myosin) have been available for some time. These structures show that significant changes in the switch motifs help to communicate the hydrolysis of ATP or GTP to the rest of the molecule, switching off G proteins and driving the movement of myosin. The structures obtained by Kikkawa *et al.*¹ reveal that equivalent structural permutations are linked to ATP hydrolysis in the switch I and switch II regions of kinesin (Fig. 1). Although not entirely unexpected, it is reassuring to see this expectation fulfilled in actual atomic structures.

What are the consequences of the ATP-dependent structural changes in the switch regions of kinesin? The most dramatic change occurs in a subdomain of the kinesin head. This subdomain, which the authors refer to as the 'switch II cluster', is linked to switch II, and consists of two structural features called α -helices, and three loops. One of the α -helices, $\alpha 4$ (sometimes referred to as the switch II helix), is a prominent component of the microtubule-binding face of the kinesin head⁵. This helix undergoes nucleotide-dependent changes in length — when the kinesin head is bound to ADP rather than the ATP analogue, the helix is longer by about two turns, at the expense of one of the loops, which shortens. The helix also shows nucleotide-dependent rotation of about 20° relative to the rest of the kinesin head (Fig. 1), causing other elements of the switch II cluster to follow this movement. This rotation, in turn, is communicated to the kinesin neck domains, which previous studies have shown to have a decisive role in the generation of force and movement by kinesin⁶.

These crystal structures were of kinesin alone, not attached to microtubules. With just these structures available, Kikkawa *et al.* might have proposed that the generation of force by kinesin involves this slight rotation of the $\alpha 4$ helix relative to the rest of the motor domain and the microtubule. But, because they also obtained cryo-electron-microscopic images of microtubule-bound kinesin heads in both activity states, a different picture emerges. Such images show low-resolution, three-dimensional contours of the kinesin head in contact with the microtubule surface. The atomic structures can be

Molecular motors

Switching on kinesin

Manfred Schliwa and Günther Woehlke

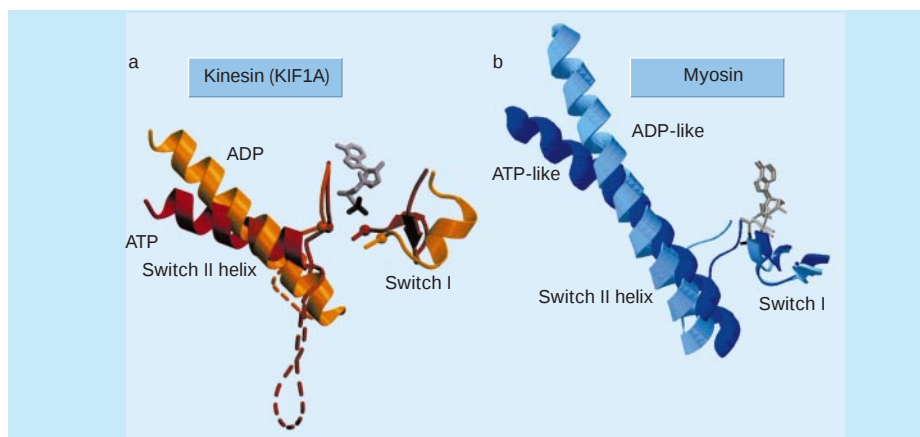
A crystal structure helps us to understand how an enzyme works. Better yet are crystal structures of the enzyme in different states of activity, which have now revealed the intricate workings of a molecular motor.

Twist-off caps are an effective way of sealing bottles. A small rotation of the cap in one direction tightens it, while a small turn in the opposite direction removes it from the bottle. A twist-off cap seems an unlikely model for a molecular motor protein, but, on page 439 of this issue¹, Kikkawa and colleagues suggest that a roughly 20° rotation of the motor region of kinesin helps it to tighten its grip on the surface of the filamentous tracks it moves along. This finding was made possible by the crystallization of a motor region in two functionally significant states — one bound to the nucleotide adenosine diphosphate (ADP), and the other in complex with an analogue of adenosine triphosphate (ATP). The 'ATP-like' conformation had, until now, resisted all attempts

to crystallize it. The authors' success ushers in a new era in the study of kinesin's mechanics and chemistry.

Kinesin motors move in one direction along linear tracks called microtubules by virtue of cyclical changes in the conformation (shape) of kinesin, driven by the hydrolysis of one molecule of ATP (thereby producing ADP) per step. They share this feature with a different type of motor, myosin, which uses actin filaments as a track. These two classes of motor have no apparent similarities in their amino-acid sequences. But their atomic structures, obtained by X-ray crystallography, reveal a remarkable conformational similarity between the central portion of myosin, which includes the nucleotide-binding site, and the 'head'

Figure 1 Comparison of the switch I and switch II regions of two molecular motor proteins. a, A kinesin protein (KIF1A; based on Fig. 1c, page 440) and b, myosin from *Dictyostelium*. a, Orange denotes KIF1A bound to ADP; red denotes KIF1A bound to AMPPCP (a non-hydrolysable analogue of ATP), as determined by Kikkawa *et al.*¹. The dashed line represents one of the loops in the switch II cluster. b, The region near the nucleotide-binding site in myosin. Light blue denotes the ADP-like state (myosin bound to MgADP-BeFx (ref. 7); structure extracted from the RCSB Protein Data Base, accession number 1MMD). The long, straight switch II helix here is bent and slightly shifted in the ATP-like state (myosin bound to MgADP-vanadate⁸, dark blue; accession number 1VOM). The two structures of each motor are aligned on the nucleotide-binding loop (omitted for clarity), so that the nucleotide (grey) can serve as a reference for



the structural changes in switch I and switch II. The small rotation of the switch II helix in myosin resembles that of the analogous structure in KIF1A, even though the translation of this conformational change into movement differs for these two motors⁹.

fitted ('docked') into these contours to determine whether the structures obtained by crystallization correspond to actual conformations of kinesin heads on microtubules. These docking studies add an interesting twist: they suggest that the switch II cluster stays fixed in one position, while the bulk of the motor domain rotates by about 20° relative to the microtubule. The authors propose that this 'screw-cap' rotation tightens the grip of the motor domain on the microtubule surface when ATP occupies the nucleotide-binding site, presumably because the $\alpha 4$ helix slides a tiny bit deeper into a complementary groove in the microtubule surface.

Not everybody in the kinesin field is likely to agree with these docking studies — they involve a certain degree of eye-balling and rarely produce a perfect match. But the authors offer a testable working model. In their view, the 20° twist also produces enough displacement of the motor tip towards one end of the microtubule to bias movement in that direction. It remains to be seen whether this displacement (just a few ångströms) does indeed generate sufficient directional bias. Kikkawa *et al.* also suggest that their findings, obtained using a monomeric, and therefore somewhat unusual, kinesin, can be extended to other proteins in the kinesin family. A comparison of the monomeric kinesin with crystal structures of other kinesins suggests that this is plausible, but confirmation is needed.

Kikkawa *et al.*'s proposal of a twist-on ATP-bound state and a twist-off ADP-bound state for kinesin, allowing the motor to alternate between tight and weak binding conformations, is likely to stimulate much discussion. It certainly is easy to remember. Next time you untwist the cap on a bottle of your favourite drink, think of kinesin. ■

Manfred Schliwa and Günther Woehlke are at the Institute for Cell Biology, University of Munich, Schillerstrasse 42, 80336 Munich, Germany.

e-mail: schliwa@bio.med.uni-muenchen.de

1. Kikkawa, M. *et al.* *Nature* **411**, 439–445 (2001).
2. Kull, F. J. *et al.* *Nature* **380**, 550–555 (1996).
3. Sablin, E. P. *et al.* *Nature* **380**, 555–559 (1996).
4. Vale, R. D. *J. Cell Biol.* **135**, 291–302 (1996).
5. Woehlke, G. *et al.* *Cell* **90**, 207–216 (1997).

6. Rice, S. *et al.* *Nature* **402**, 778–784 (1999).
7. Fisher, A. J. *et al.* *Biochemistry* **34**, 8960–8972 (1995).
8. Smith, C. A. & Rayment, I. *Biochemistry* **35**, 5404–5417 (1996).
9. Woehlke, G. & Schliwa, M. *Nature Rev. Mol. Cell Biol.* **1**, 50–58 (2000).

Bose–Einstein condensation

Getting excited about helium

Randall G. Hulet

Creating a quantum fluid from a gas of excited helium atoms is not easy — the atoms tend to self-destruct. But two groups in France have pulled it off.

When liquid helium is cooled to a temperature of about two degrees above absolute zero (2 kelvin), something strange and spectacular happens: it suddenly acquires the ability to flow without friction. This transformation, first observed in the 1930s and known as Bose–Einstein condensation (BEC), occurs when the helium atoms join together into a single quantum state (some of the history is recounted in the obituary of Jack Allen on page 436). The collective behaviour of BEC is responsible not only for the 'super' in these liquid-helium superfluids, but also, in an indirect sense, for superconductivity in solids. Now helium is once again making BEC news, thanks to breakthroughs by two groups in France. Writing in *Science* and *Physical Review Letters*, Robert *et al.*¹ and Pereira Dos Santos *et al.*² report the creation of Bose–Einstein condensates of helium, but this time in a gas of excited atoms.

Although helium now joins rubidium³, lithium⁴, sodium⁵ and hydrogen⁶ as elements that have been turned into condensates in the gas phase, the new work is not just another step through the periodic table. The results are exciting because, for the first time, the helium atoms are not in the lowest

electronic state, the ground state, when BEC occurs. Instead, they are 'metastable' atoms that have nearly 20 electron volts (eV) of internal energy. This excitation energy is huge compared with the thermal energy of a supercooled condensate (just 10^{-10} eV per atom).

Such a large internal energy offers a route to detecting single atoms, but it also posed a serious challenge to achieving BEC in the first place. As shown in Fig. 1a (overleaf), the collision of two metastable helium atoms (He^*) can result in the ionization of one of the He^* atoms, leaving the other in the ground state. Because the metastable state has such a large excitation energy, the rate of ion production (Penning ionization) is extraordinarily large — so large, in fact, that the gas of metastable atoms destroys itself in just a few milliseconds.

The success of the new experiments hinges on a bold prediction by Shlyapnikov, Walraven and collaborators⁷ that the rate of Penning ionization could be reduced by several orders of magnitude, perhaps by as much as a factor of 10,000, if the He^* atoms were spin polarized. Particles like atoms and electrons have a quantum-mechanical spin and can be thought of as tiny bar

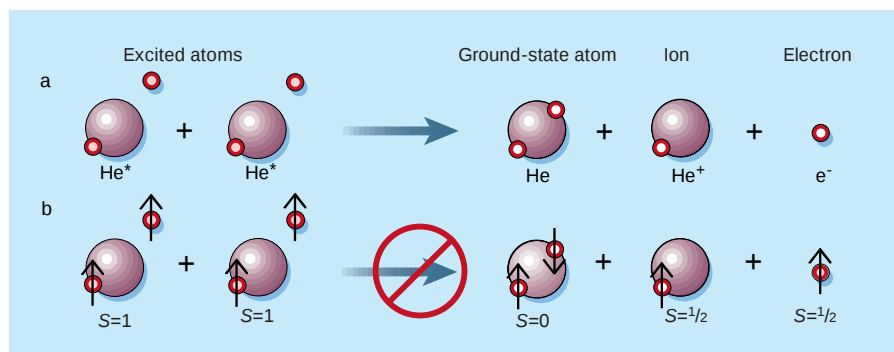


Figure 1 Collisions between excited helium atoms. a, The collision of two unpolarized metastable helium atoms (He^*) results in Penning ionization, producing a ground-state helium atom, a helium ion (He^+) and a free electron. The rate for this process is so large that it would prevent Bose–Einstein condensation. b, Penning ionization can be suppressed by spin polarizing the atomic electrons. Because spin must be conserved, two polarized He^* atoms are unable to undergo Penning ionization. The initial atoms have total spin, $S = 2$ (each of the four electrons contributes $1/2$), whereas the products can have spin no greater than 1. Two groups^{1,2} have now used this process to achieve Bose–Einstein condensation in a metastable helium gas.

magnets. Each electron has a spin, S , of $1/2$ and can point either up or down. When helium atoms are spin polarized, the total spin of two He^* atoms is the sum of the spins of their four electrons ($S = 2$). After Penning ionization, the total spin of the products — a ground-state helium atom with zero spin, a He^+ ion and a free electron each with spin $1/2$ — can be no more than 1 (Fig. 1b). Because spin must be conserved in quantum mechanics, Penning ionization is severely suppressed. Spurred on by this prediction, several groups began experiments to make BEC in spin-polarized He^* .

For BEC to occur, the quantum wavelengths of the atoms must be greater than their average separation, which for liquid helium is about 0.4 nanometres. For experiments on liquid helium, this condition is met only at temperatures below 2.2 K. But for helium gas, in which the density is roughly a billion times lower, ultralow temperatures of around $1\ \mu\text{K}$ are needed. Both groups in France achieve this by using what are now routine methods of laser cooling, magnetic trapping and forced evaporative cooling. Magnetic traps confine magnetic atoms, such as He^* , in bowl-shaped magnetic fields. A bonus of this technique is that the magnetically trapped atoms are necessarily spin polarized, and hence meet the condition required for suppressing Penning ionization. A new challenge for these BEC experiments is producing metastable helium. Because no lasers operate at the extreme ultraviolet wavelength of 63 nanometres required to resonantly excite ground-state helium, the researchers had to rely on an electrical discharge to provide the necessary excitation energy. This process is woefully inefficient, converting only one in every 100,000 atoms.

Although similar in every other respect, the two new experiments differ significantly in their methods for detecting BEC. A sig-

nature of BEC is the sudden appearance of atoms moving very slowly. Whereas Pereira Dos Santos *et al.*² used optical absorption to image the He^* , a common technique in alkali-metal BEC experiments, Robert *et al.*¹ exploited the 20 eV of excitation energy released when He^* hits a surface. When the He^* atom falls on a charged-particle detector it is readily ionized and initiates a shower of electrons that are easily amplified and detected as an electrical pulse. This high detection efficiency is one of the things that makes a condensate of He^* atoms different from the alkali condensates or liquid helium.

At a sufficiently low temperature of about $5\ \mu\text{K}$, and with 10^5 to 10^6 atoms in their traps, both teams observed the characteristic low-velocity signal of BEC, confirming that Penning ionization is suppressed in a gas of spin-polarized He^* . Pereira Dos Santos *et al.*² report the suppression to be a factor of at least 2,000, in agreement with theory. Both groups also find that the elastic scattering cross-section, an important quantity governing the rate at which atoms collide and come into thermal equilibrium, is especially large for He^* . It was the fortunate combination of a high rate for achieving thermal equilibrium and a low rate of Penning ionization that made these experiments successful.

What's next for helium condensates? The ability to detect He^* atoms with almost perfect efficiency clears the way for some intriguing applications. One promising direction to explore is what happens to BEC, which by definition is a many-body phenomenon, when there are just a few atoms. Many questions regarding the fluctuations of condensate number and quantum phase may be addressed. A second avenue is to use He^* for atom lithography, in which nanoscale features are drawn on a surface by directly depositing atoms. Previous experiments have exposed a surface to beams of He^* to deposit features as fine as a few tens



100 YEARS AGO

Mr. Andrew Carnegie, the American millionaire, has come forward with a proposal to provide free University education to the youth, both male and female, of Scotland, and offers to place the sum of two millions of pounds in the hands of trustees... There can be but one opinion regarding the large-heartedness which prompts so magnificent a benefaction, and the whole nation will hope that a sound result may be obtained through so noble a gift... Two obvious criticisms evoked by the bare statement that has been made public may, without detracting from the generous intention of the donor, be noted. In the first place, the consequence of the gift as adumbrated must be that secondary education will, in Scotland, alone be unendowed... Secondly, the gift is no endowment of the Scottish Universities, but it may, on the contrary, be an embarrassment to them. It means the creation of some sixteen hundred bursaries, each of the value of nine pounds, in each of the Universities. This will not bring an influx of sixteen hundred students to each University, but, if Mr. Carnegie's intention be realised, we take it there be a considerable increase in the number—sufficient, indeed, to swamp the existing equipment for teaching.

From *Nature* 23 May 1901.

50 YEARS AGO

In the course of experiments on the persistence of deposits from aqueous suspensions of different particle sizes of volatile insecticides, it has been found that the residual toxicity of particles of any one size is influenced considerably by the type of material to which they are applied. Most striking results have been obtained on mud blocks made from 'murrum', a lateritic ironstone, used in the construction of walls of houses in Uganda. Crystals of all insecticides used rapidly disappear from the surface of these blocks when they are kept at 78°F (25°C), and even those of DDT, which is usually regarded as a contact insecticide with a long residual life, are no longer visible after only a few days... When DDT and 'Dieldrin' crystals are no longer visible on the surface, the mud blocks lose their toxicity to mosquitoes (*Aedes aegypti*, L.) exposed to them for long contact periods.

From *Nature* 26 May 1951.

of nanometres^{8,9}. With metastable atoms it may be possible to create more complex patterns using atom holography¹⁰. The efficiency of these processes, although currently low, could be enhanced significantly with an efficient 'atom laser' constructed from a metastable BEC.

So although Bose–Einstein condensates of He* and liquid helium are composed of the same atomic element, they have very different characteristics. The unique properties of a BEC of metastable helium gas will give researchers a powerful tool with which to learn more about this fascinating state of matter.

Randall G. Hulet is in the Department of Physics

and Astronomy and Rice Quantum Institute, Rice University, Houston, Texas 77251, USA. e-mail: randy@atomcool.rice.edu

1. Robert, A. *et al.* *Science* **292**, 461–464 (2001).
2. Pereira Dos Santos, F. *et al.* *Phys. Rev. Lett.* **86**, 3459–3462 (2001).
3. Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. *Science* **269**, 198–201 (1995).
4. Bradley, C. C., Sackett, C. A., Tollett, J. J. & Hulet, R. G. *Phys. Rev. Lett.* **75**, 1687–1690 (1995).
5. Davis, K. B. *et al.* *Phys. Rev. Lett.* **75**, 3969–3972 (1995).
6. Fried, D. G. *et al.* *Phys. Rev. Lett.* **81**, 3811–3814 (1998).
7. Shlyapnikov, G. V., Walraven, J. T. M., Rahmanov, U. M. & Reynolds, M. W. *Phys. Rev. Lett.* **73**, 3247–3250 (1994).
8. Nowak, S., Pfau, T. & Mlynek, J. *Appl. Phys. B* **63**, 203–205 (1996).
9. Bard, A. *et al.* *J. Vacuum Sci. Technol. B* **15**, 1805–1810 (1997).
10. Morinaga, M. *et al.* *Phys. Rev. Lett.* **77**, 802–805 (1996).

RNA interference

The short answer

Brenda L. Bass

One way of seeing what a gene does is to block its messenger RNA and note the effects. New work should make the approach more broadly applicable.

RNA interference (RNAi) was discovered only a few years ago¹, but many scientists find it hard to imagine life without it. Once the sequence of a gene is known, RNAi offers a quick and easy way to determine its function, and the technique is accessible to a scientist in a small lab, as well as to a consortium attempting to assign function to the genes of an entire chromosome^{2,3}. But although RNAi is now routine in laboratories studying a wide range of organisms, its use in mammalian cells has been problematic. On page 494 of this issue⁴ Tuschl and colleagues describe research that paves the way for successful RNAi in mammalian cells.

The basic idea behind RNAi is shown in the right-hand part of Fig. 1 — this is the sequence-specific pathway indicated by blue arrows. A double-stranded RNA (dsRNA) matching a gene sequence is synthesized *in vitro* and introduced into a cell. The dsRNA feeds into a natural, but poorly understood, biological pathway, and is broken into short pieces called short interfering (si) RNAs⁵. With the help of cellular enzymes that have not yet been well characterized⁶, the siRNA triggers the degradation of the messenger RNA that matches its sequence. This often leads to adverse consequences for the organism, evident in an aberrant phenotype, that allow the gene's function to be identified.

RNAi was first discovered in the nematode worm *Caenorhabditis elegans*¹, but is present in many other organisms (the fruit fly *Drosophila*, certain parasitic protozoa, and plants, for instance), and so seems to represent an ancient pathway⁷. Nonetheless, researchers have always been pessimistic

about applying RNAi to mammalian cells, because exposing such cells to dsRNA, of any sequence, triggers a global shut-down of protein synthesis⁸. This nonspecific pathway is indicated on the left of Fig. 1 by a red arrow. The lore has been that this pathway would mask any sequence-specific effects that might occur from the RNAi pathway.

But it almost always pays to consider how one's own research fits in with previous observations. Tuschl and colleagues were, it seems, being especially diligent in this respect. Their earlier work showed that the

siRNA intermediates themselves could initiate RNAi, at least in non-mammalian cells⁵. However, the nonspecific pathway requires longer dsRNA, and will not occur with dsRNAs shorter than around 30 base pairs^{8–10}. They don't say as much in the paper, but one presumes that Tuschl's group began with the idea that, because of this size discrimination, siRNAs might be able to bypass the more global, nonspecific response. They turned out to be right.

First, Tuschl and colleagues⁴ tested whether siRNAs could trigger RNAi in mammalian cells, as had been observed in non-mammalian cells. They assayed the ability of siRNA to target various luciferase transgenes, for which gene expression is easily quantified by measuring luminescence. siRNAs were transfected with cationic liposomes into various mammalian tissue culture cells (NIH/3T3, COS-7, HeLa and 293 cells), as well as into a *Drosophila* cell-culture line for comparison. Indeed, the authors observed reproducible, sequence-specific siRNA inhibition in the mammalian cells, with no sign of the nonspecific effects. In contrast, with longer RNAs, luciferase expression was reduced with every dsRNA tested, no matter what its sequence. Superimposed on the nonspecific inhibition was a sequence-specific inhibition, suggesting that both pathways can operate simultaneously. (As shown in Fig. 1, the two pathways probably compete for the long dsRNA.) Importantly, Tuschl and co-workers went on to show that siRNAs are not only effective at targeting the transgene luciferase, but also at targeting naturally occurring, endogenous genes.

Of course, the story is not as neat and tidy as I have described it here. As the authors are

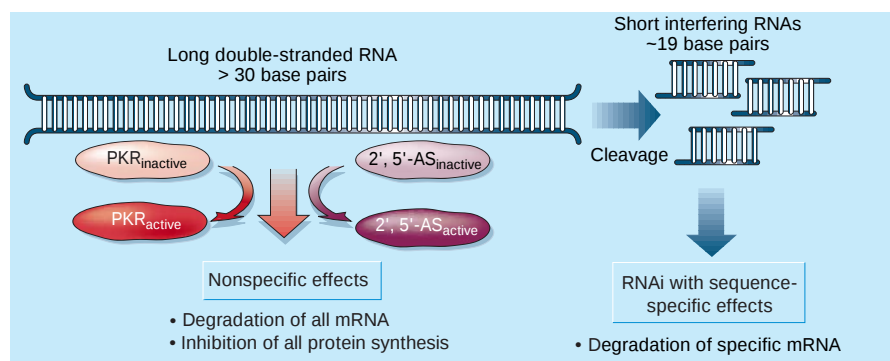


Figure 1 Mammalian cells have at least two pathways that compete for double-stranded RNA (dsRNA). In the RNAi, or sequence-specific, pathway (blue arrows), the initiating dsRNA is first broken into short interfering (si) RNAs. siRNAs have sense and antisense strands of about 21 nucleotides that form 19 base pairs to leave overhangs of two nucleotides at each 3' end. siRNAs are thought to provide the sequence information that allows a specific messenger RNA to be targeted for degradation. The nonspecific pathway (red arrow) is triggered by dsRNA of any sequence, as long as it is at least 30 base pairs long. The nonspecific effects occur because dsRNA activates two enzymes: PKR, which in its active form phosphorylates the translation initiation factor eIF2a to shut down all protein synthesis, and 2', 5' oligoadenylate synthetase (2', 5'-AS), which synthesizes a molecule that activates RNase L, a nonspecific enzyme that targets all mRNAs. The nonspecific pathway represents a host response to stress or viral infection; in the second case, the activating dsRNA is thought to derive from viral replication.

careful to point out, inhibition by siRNAs is effective in mammalian cells, but gene expression is not eliminated completely as it is in *Drosophila* cells. Further, siRNA techniques in mammalian cells have some of the same drawbacks associated with antisense RNA, another technique used to prevent expression of particular genes. In both cases, success depends on the cell type, as well as on the level of expression of the gene to be targeted. That apart, however, RNAi has repeatedly proven itself to be more robust than antisense techniques: it works more often, and typically decreases expression of a gene to lower levels, or eliminates it entirely. And, as Tuschl and colleagues show, even in mammalian cells, siRNAs are effective at concentrations that are several orders of magnitude below the concentrations typically used in antisense experiments.

One of the most important aspects of the new work is the further research it will inspire. Although RNAi works in mouse eggs and embryos^{11,12}, scientists have been reluctant to invest time in applying it to other mammalian cells because of reported problems^{13,14}. Now we will see studies aimed at optimizing the use of siRNAs, as well as at understanding why conventional RNAi, with longer dsRNA, works in eggs and embryos. Might these cells lack the non-specific pathway?

The RNAi technique has had a huge impact in studies of non-mammalian systems. Use of siRNA in mammalian cells could be just as far-reaching, with the applications extending to functional genomics and therapeutics. But various technical issues must be addressed, especially for large-scale applications. For instance, dsRNA can be delivered to *C. elegans* by feeding or soaking, but effective delivery of siRNAs to mammalian cells will not be so simple. The analysis of *C. elegans* phenotypes is aided by short generation times and a wealth of information about the worm's morphology and behaviour; developing rapid ways to screen mammalian cells, or whole organisms, will take some time and thought.

So far I have discussed RNAi as a technique. But of course the pathway does not exist in cells solely to make life easier for scientists. RNAi is a natural biological pathway, albeit one we don't quite understand yet. Especially for those with a long-standing interest in the roles of dsRNA, Tuschl and colleagues' paper is interwoven with information about how RNAi coexists with previously characterized dsRNA pathways. This is especially interesting because dsRNA-binding proteins are usually not sequence specific and will bind any dsRNA. A single dsRNA can interact with proteins of different pathways so that the pathways compete. The different ratios of specific to nonspecific inhibition observed by the authors are probably telling us something about the particu-

lar constellation of dsRNA-binding proteins in the different cell types and how they compete with RNAi. Regardless of that, the new study shows that one way dsRNA pathways can coexist is to require different lengths of dsRNA. This is good news for cells — and for researchers.

Brenda L. Bass is in the Department of Biochemistry and Howard Hughes Medical Institute, University of Utah, 50 North Medical Drive, Room 211, Salt Lake City, Utah 84132, USA. e-mail: bbass@howard.genetics.utah.edu

1. Fire, A. *et al.* *Nature* **391**, 806–811 (1998).
2. Barstead, R. *Curr. Opin. Chem. Biol.* **5**, 63–66 (2001).
3. Kim, S. K. *Curr. Biol.* **11**, R85–R87 (2001).
4. Elbashir, S. M. *et al.* *Nature* **411**, 494–498 (2001).

5. Elbashir, S. M., Lendeckel, W. & Tuschl, T. *Genes Dev.* **15**, 188–200 (2001).
6. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. *Nature* **409**, 363–366 (2001).
7. Sharp, P. A. *Genes Dev.* **15**, 485–490 (2001).
8. Hunter, T., Hunt, T., Jackson, R. J. & Robertson, H. D. *J. Biol. Chem.* **250**, 409–417 (1975).
9. Manche, L., Green, S. R., Schmedt, C. & Mathews, M. B. *Mol. Cell Biol.* **12**, 5238–5248 (1992).
10. Minks, M. A., West, D. K., Benveniste, S. & Baglioni, C. *J. Biol. Chem.* **254**, 10180–10183 (1979).
11. Wianny, F. & Zernicka-Goetz, M. *Nature Cell Biol.* **2**, 70–75 (2000).
12. Svoboda, P., Stein, P., Hayashi, H. & Schultz, R. M. *Development* **127**, 4147–4156 (2000).
13. Ui-Tei, K., Zanno, S., Miyata, Y. & Saigo, K. *FEBS Lett.* **479**, 79–82 (2000).
14. Caplen, N. J., Fleenor, J., Fire, A. & Morgan, R. A. *Gene* **252**, 95–105 (2000).

Chemistry

Synthetic lessons from quinine

Steven M. Weinreb

As the oldest, naturally occurring, treatment for malaria, quinine has been a target for synthetic chemistry for 150 years. At last, modern techniques provide full control over the synthetic molecule.

Nowadays the isolation and structure determination of a new biologically significant, naturally occurring molecule stirs the interest of synthetic chemists as well as biologists. In many cases, such a discovery is quickly followed by a laboratory synthesis. Organic chemists now have the ability and tools to synthesize virtually any large and complex natural product, but there are a few 'classical' compounds that have been revisited by the synthetic community time and time again. One such compound is the alkaloid quinine, chiefly used to treat malaria. In the *Journal of the American Chemical Society*, Gilbert Stork and co-workers¹ show that chemical 'firsts' are still possible — even when dealing with a natural product that has been used in medicine for over 300 years, and which has been a prime target for synthetic organic chemists for 150 years.

The Countess of Chinchon, the consort of the Spanish Viceroy of Peru, first popularized quinine as a treatment for malaria in the early seventeenth century. Natural quinine and related alkaloids are extracted from the bark of *Cinchona* trees. These alkaloids were first isolated in pure form in the early nineteenth century, but the structures of the major compounds were not worked out for another 100 years. Like many drugs, quinine ($C_{20}H_{24}N_2O_2$) can exist in several structural forms. It contains 'asymmetric' carbon atoms, which are connected to four different chemical groups and so allow the overall structure to be arranged in different ways. For quinine, 16 such stereoisomers are possible, but only one corresponds to the active form of the drug (Fig. 1, overleaf).

Because the supply of quinine depended

on the political vagaries of the producing countries — especially during wartime — organic chemists soon became interested in synthesis as a way to ensure a constant supply of the drug. Early attempts to synthesize quinine are legendary. In 1856, even before the structure of quinine was elucidated, William Henry Perkin² naively attempted to prepare the alkaloid by oxidizing allyltoluidine ($C_{10}H_{13}N$). In the course of this work he inadvertently produced the first synthetic dye, which was the main event responsible for the founding of the organic chemical industry.

When detailed structural information on the *Cinchona* alkaloids became available in the twentieth century, more rational approaches to quinine synthesis were devised. In 1944, two Harvard chemists, R. B. Woodward and William Doering, announced the first formal synthesis of quinine³, despite the limited synthetic and spectroscopic tools available at that time. Although they did not actually make quinine itself, they succeeded in synthesizing an intermediate compound known as homomeroquinene (Fig. 1). This compound is a degradation product of quinine, and had reportedly been converted back to the alkaloid by Paul Rabe in 1918 (although that account is now in dispute)^{1,4}. So Woodward and Doering had developed a synthetic route to quinine insofar as they assumed that Rabe's part of the sequence would work. But their synthesis of homomeroquinene was inefficient, largely because they could not control the spatial arrangement of atoms around the asymmetric carbons. This meant they actually generated several stereoisomers of homomeroquinene, which then

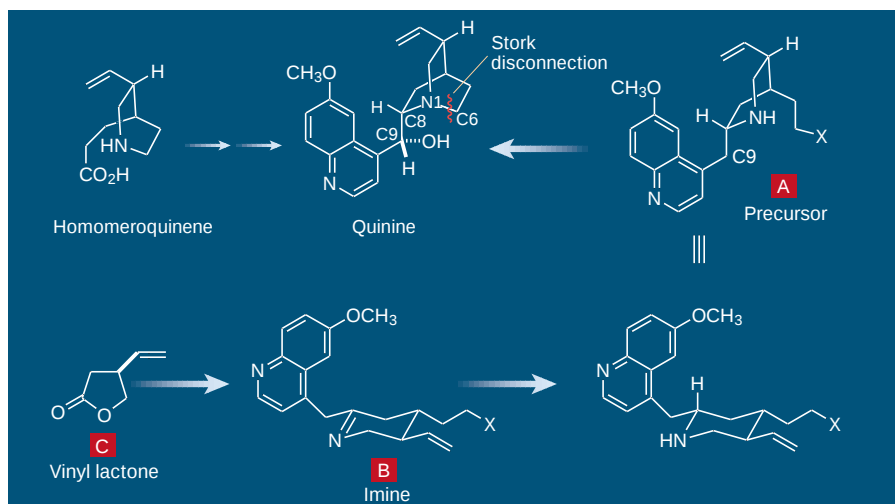


Figure 1 Routes to synthetic quinine — the oldest known treatment for malaria. A famous wartime synthesis of quinine was claimed by Woodward and Doering³ in 1944, in response to limited supplies of the natural drug. But they achieved only a partial synthesis of the compound — making the intermediate homomeroquinene and relying on a previously reported sequence for turning this into the final product. Later work⁵ produced a full synthesis of quinine, but it could not control the crucial arrangement of atoms around two central carbons (C8 and C9). This meant that several structural variations (stereoisomers) of quinine were produced, which then had to be separated from the active drug. Stork and colleagues¹ have now succeeded in synthesizing quinine by a different route, which involves a more complex intermediate (A), but which yields the correct configuration at C8 and C9.

had to be put through a tedious separation process.

Researchers got closer to a full stereoselective synthesis of quinine in the 1970s, when Milan Uskoković and co-workers at Hoffmann-La Roche developed a considerably more efficient strategy that controlled most — but still not all — configurations of the asymmetric carbons in the molecule^{5,6}. The Uskoković approach was similar to that of Woodward and Doering in the sense that a homomeroquinene derivative was a pivotal intermediate. The Hoffmann-La Roche synthesis produced equal amounts of quinine and its stereoisomer quinidine, which differs from quinine in the arrangement of the bonds around two carbon atoms (C8 and C9 in Fig. 1). Quinidine is also a naturally occurring *Cinchona* alkaloid, and is sometimes used to treat heart conditions.

Normally, organic chemists design the synthesis of a complex molecule by working backwards from the final product in order to come up with intermediates that are simpler and easier to make⁷. In other words, one works in a 'retrosynthetic' direction, conceptually breaking key bonds (disconnection) with the intention of forming those strategic bonds in the laboratory. Before Stork's work¹, quinine syntheses had all been based on a strategy involving the key retrosynthetic disconnection of the bond between C8 and N1, effectively making homomeroquinene the intermediate (Fig. 1). Stork, an acclaimed master of organic synthesis who has worked intermittently on quinine for over 55 years, recognized that disconnecting the C6–N1 bond in the alkaloid could offer

complete stereochemical control over the intermediate and final products.

Why was this approach not previously investigated? At first glance, this disconnection produces an intermediate compound (A in Fig. 1) that is almost as complex as quinine itself, and so would be difficult to make. Although this conclusion may have been valid several years ago, by using modern synthetic methodology, Stork has designed a direct and efficient route to A by way of a stereoselective reduction of the imine (B). In turn, B can be prepared in a pure isomeric

form in about ten simple steps from the readily available vinyl lactone (C). In the final step, the necessary functionality and stereochemistry is cleanly introduced at C9 by an auto-oxidation.

This last step was based on an important observation previously made by the Uskoković group⁵. None of the steps in the Stork synthesis is completely new, but their arrangement is what makes this an elegant and efficient strategy. In addition, the Stork paper is written with an insight and historical perspective (as well as correcting some myths) rarely seen in the primary chemical literature, and should be required reading for all students of organic chemistry.

Although quinine is still used worldwide to treat malaria, it is slowly being supplanted by newer drugs. Along with the fact that the alkaloid is still readily obtained from natural sources, it is unlikely that total synthesis of quinine will ever provide an economical supply of the drug. Nonetheless, the lessons in synthetic methodology and strategy learned over the years from the many chemists who have worked on quinine have been significant, and it seems unlikely that the Stork synthesis will be the final word in this area.

Steven M. Weinreb is in the Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802, USA.
e-mail: smw@chem.psu.edu

1. Stork, G. *et al.* *J. Am. Chem. Soc.* 123, 3239–3242 (2001).
2. Perkin, W. H. *J. Chem. Soc.* 69, 596–637 (1896).
3. Woodward, R. B. & Doering, W. E. *J. Am. Chem. Soc.* 66, 849 (1944).
4. Rouhi, A. M. *Chem. Eng. News* 7 May, 54–56 (2000).
5. Grethe, G., Lee, H. S., Mitt, T. & Uskoković, M. R. *J. Am. Chem. Soc.* 100, 589–593 (1978).
6. Verpoorte, R., Schripsema, J. & van der Leer, T. in *The Alkaloids* Vol. 34, Chap. 6 (ed. Brossi, A.) 331–398 (Academic, San Diego, 1988).
7. Corey, E. J. & Cheng, X.-M. *The Logic of Chemical Synthesis* (Wiley, New York, 1989).

Carbon cycle

Fertile forest experiments

Eric A. Davidson and Adam I. Hirsch

Long-term experiments under realistic conditions are beginning to deliver data on how forests — or at least some forests — will react to increasing levels of CO₂ in the atmosphere.

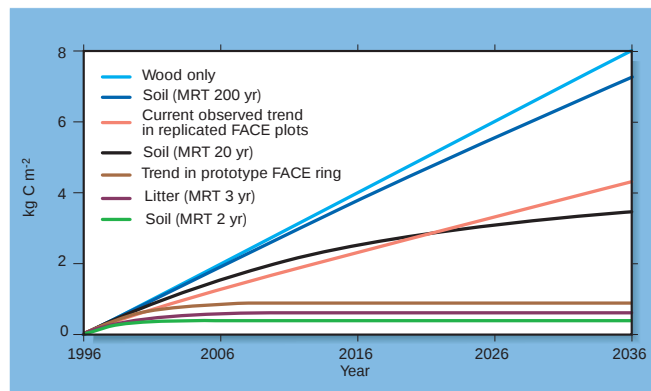
The global experiment of increasing atmospheric CO₂ concentrations by burning fossil fuels has neither a control nor replicates. So it is difficult to quantify how much faster the world's forests might be growing under high CO₂ conditions. Higher levels of CO₂ can clearly make plants grow better. But will Earth's vegetation absorb from the atmosphere, and retain, much of the CO₂ pouring out of our exhaust pipes and smoke stacks? If it does, then the threat of global warming from increasing CO₂ would be less severe. Current estimates

of annual terrestrial plant uptake of carbon attributable to 'CO₂ fertilization' are within the range 0.5–2.0 × 10¹⁵ g, which is about 8–33% of annual fossil-fuel emissions¹. What are needed of course, are controlled, replicated experiments under realistic field conditions to help reduce the uncertainty of these figures.

Papers in this issue by Oren *et al.*² and Schlesinger and Lichter³ describe just that — well-designed studies to test the effects of CO₂ fertilization, in these cases at the scale of a forest stand in the United States. The data

Box 1 Carbon sequestration at FACE

The figure here shows projected cumulative carbon sequestration attributed to CO₂ fertilization at the Duke 'Free air CO₂ enrichment' site, with different assumptions about carbon allocation, mean residence times (MRT), and growth limitation by nutrients and water. The untreated forest has been accumulating carbon at a rate of about 0.3 kg C m⁻² yr⁻¹, which is common for loblolly pine plantations in the southeastern United States¹². Here we show the large differences in the projected cumulative carbon sink when an extra 0.2 kg C m⁻² yr⁻¹ from CO₂ fertilization⁷ is allocated entirely to wood, to leaf litter, or to one of three soil pools with residence times of 2, 20 and 200 years. Wood, and soil fractions with long



residence times, are expected to sequester large amounts of carbon during the next few decades. But carbon inputs to litter and to the rapidly cycling soil pools result in rapid re-equilibration at a new steady state, with a relatively trivial

cumulative sequestration. The rate of carbon accumulation is calculated from

$$C_t = C_0 e^{-kt} + (I/k)(1 - e^{-kt})$$

where C_0 is the stock at the

beginning of the experiment (1996), C_t is the stock after t years, I is the carbon input and k is the decomposition constant ($k = 1/\text{MRT}$). The initial 1996 carbon stocks, and inputs of wood, litter layer and soil, and the residence times of the litter layer, are from Schlesinger and Lichter³. Projections are also shown for the recent trends in the replicated FACE plots, where about half of the additional inputs from CO₂ fertilization were allocated to wood and about half to rapidly cycling litter and soil pools⁷. In contrast, if the increased wood allocation ceases as it did in the prototype FACE ring, the extra carbon sink will level off sooner at a much lower cumulative sum. E.A.D. & A.I.H.

show only modest carbon storage in the soil and leaf-litter pools, and also provide preliminary evidence that shortage of nutrients and water may limit the long-term response of trees to increased levels of CO₂.

Well-watered and fertilized citrus trees clearly grow better when exposed to high CO₂ (ref. 4). But the effect on forest trees is uncertain because most forests have limited supplies of other resources needed for growth. Equally important is how trees allocate their photosynthetic output. As described in Box 1, if the increased carbohydrate is used to produce more wood, then that carbon is likely to reside in the forest for many decades. But if it goes instead to leaves, which die and decompose within a few years, the carbon will quickly be released back to the atmosphere. Allocation of carbon to roots and associated fungi could create either a short- or long-term sink, depending on whether the carbon is stabilized in the soil or is rapidly decomposed to CO₂ after the roots die⁵.

Both papers^{2,3} describe results from the 'Free air CO₂ enrichment' (FACE) experiment at the Duke forest of Duke University, North Carolina. The experiment began with the construction of a prototype 700-m² ring-shaped fumigation structure in a young loblolly pine plantation in 1993. This was followed by construction of three treatment and three control rings in 1996. The treatment rings are fumigated with high-CO₂ air to maintain concentrations 200 µl l⁻¹ above ambient; control plots are fumigated with ambient air (CO₂ concentration 365 µl l⁻¹). Previous CO₂ enrichment experiments used chambers containing at most a few saplings, but each of the Duke FACE rings con-

tains several dozen trees that are now about 15 m tall.

Oren *et al.*² (page 469) report enhanced wood growth in the prototype FACE ring during the first three years of CO₂ fumigation, but then a return to pre-fumigation growth rates. To see why increased growth stopped, they fertilized half of the ring with nitrogen while maintaining the high-CO₂ treatment. Trees on the nitrogen-fertilized half resumed increased growth, while those on the other half did not. Outside the FACE ring, nitrogen fertilization also increased tree growth relative to controls, but the combination of nitrogen and CO₂ fertilization caused a larger response than either treatment alone. This synergistic nitrogen-CO₂ response was greater in a rainy year than a drought year, although possible time lags between drought and its effect on wood growth complicate this interpretation. These results are the clearest evidence yet that nutrients, and perhaps water, may constrain the trees' response to CO₂ fertilization in real forests.

Meanwhile, Schlesinger and Lichter³ (page 466) report on the fate of the leaf and root litter in the replicated FACE rings. The amount of forest-floor carbon has increased significantly. But given the rapid turnover of this pool (mean residence time is only about three years), the authors calculate that a new steady state will soon be established with only a slightly higher mass of litter. They also found no significant increases in mineral soil carbon attributable to higher concentrations of CO₂, supporting other evidence that increased soil CO₂ production is balancing larger carbon input into the soil⁶.

If the current trends⁷ persist in the repli-

cated FACE rings, an extra 4.3 kg C m⁻² would be sequestered after 40 years because of CO₂ fertilization (Box 1), almost entirely in wood. But if nutrients limit increased wood growth to only three years, as apparently happened in the prototype FACE ring, then this extra carbon sequestration will cease in at most ten years, at a cumulative sum of less than 1 kg C m⁻².

Human activities are also increasing the deposition of nitrogen, a component of 'acid rain', and this might lead to the synergism between nitrogen and CO₂ fertilization reported by Oren *et al.*². But this response is also complex. Soils immobilize most of the nitrogen inputs⁸, and large amounts of extra nitrogen can create nutrient imbalances and soil acidification that reduce plant growth⁹. There was little evidence of either nitrogen or CO₂ fertilization in an analysis¹⁰ of the long-term forest inventory records of the US Forest Service, with nearly all of the forest growth being attributed to tree recovery from previous forest harvesting or following cessation of agriculture. Inventory analyses address larger spatial and temporal scales than experiments such as FACE, and both approaches will be needed to disentangle the complex interactions affecting carbon storage.

Obviously, this fast-growing pine plantation in North Carolina does not represent all forests of the world. Other species grow and allocate carbon at different rates, and soils with different mineralogy and texture, and under different climatic conditions, might sequester more or less carbon¹¹ than reported here^{2,3}. Nevertheless, we now have some hard data from well-designed experiments in a realistic forest setting. Information such

as this will be essential in using models to make global extrapolations of the effects of increasing atmospheric CO₂ on the carbon cycle as a whole.

Eric A. Davidson and Adam I. Hirsch are at The Woods Hole Research Center, PO Box 296, Woods Hole, Massachusetts 02543, USA.

e-mail: edavidson@whrc.org

1. Schimel, D. et al. in *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. et al.) 65–131 (Cambridge Univ. Press, 1996).

2. Oren, R. et al. *Nature* **411**, 469–472 (2001).
3. Schlesinger, W. H. & Lichter, J. *Nature* **411**, 466–469 (2001).
4. Idso, S. & Kimball, B. *Glob. Biogeochem. Cycles* **7**, 537–555 (1993).
5. Trumbore, S. E. *Ecol. Appl.* **10**, 399–411 (2000).
6. Andrews, J. & Schlesinger, W. *Glob. Biogeochem. Cycles* **15**, 149–162 (2001).
7. DeLucia, E. et al. *Science* **284**, 1177–1179 (1999).
8. Nadelhoffer, K. J. et al. *Nature* **398**, 145–147 (1999).
9. Aber, J. et al. *BioScience* **48**, 921–934 (1998).
10. Caspersen, J. P. et al. *Science* **290**, 1148–1151 (2000).
11. Davidson, E. et al. *Nature* **408**, 789–790 (2000).
12. Schiffman, P. & Johnson, W. *Can. J. For. Res.* **19**, 69–78 (1989).

Neurobiology

Learning from a fly's memory

Randolf Menzel and Uli Müller

Two facets of learning are the formation and the retrieval of memories. Genetic manipulation of the fruitfly's brain allows them to be dissociated, and may lead to a better understanding of memory.

Seminal work done over 100 years ago^{1,2} established the basics of memory and learning. But, then as now, studies of learning were hampered by the fact that it is difficult (if not impossible) to distinguish between the formation and the retrieval of memories. Yet it will be important to make this distinction if we are to unravel fully the molecular and neural processes underlying the sequential stages of memory. On page 476 of this issue³, Dubnau and colleagues describe how they used fruitflies to tackle the problem.

During learning, memories are formed and stored in several stages, and can be retrieved at any moment when circumstances demand it. But these stages are hard to separate in laboratory studies, because memories are accessible only through retrieval. In a typical experiment, an animal might learn to associate one stimulus (such as an odour) with an electric shock. The animal would then be subjected to some sort of interfering experimental procedure, and tested some time later to see whether it can remember the association. For example, inhibitors of protein synthesis might be injected into the brain, to find out whether learning requires the synthesis of new proteins. If the interfering procedure were used some time after the association was learnt, and well before the 'retention' test, then changes in retention would be interpreted as revealing aspects of memory formation and storage, rather than retrieval. But these techniques have flaws; for example, protein-synthesis inhibitors might alter the memories that are stored.

Fruitflies (*Drosophila melanogaster*) are often used when studying complex research problems, and provide an ideal test case here. They too can learn to associate an odour with a shock, and remember the unpleasant association for a few days (this is known as

olfactory learning)⁴. Several memory phases, involving several sets of molecular machinery, are involved in acquiring and storing this information⁵.

Dubnau et al.³ now use state-of-the-art molecular genetic tools to tease apart the storage and retrieval of olfactory memories in *Drosophila*. Their target is a gene called *shibire*. This gene encodes a motor protein, dynamin, needed for the release of neurotransmitters from neurons at synapses (presynaptic function). Fruitflies in which a temperature-sensitive mutant *shibire* gene is expressed move normally at 20 °C, but become paralysed at 30 °C — the temperature at which the mutant gene is switched on. A few minutes after being shifted from the

restrictive temperature (30 °C) to the permissive one, the flies recover (as the mutant gene is switched off).

Usefully, the gene can also be targeted to particular parts of the brain, and Dubnau et al.³ use fruitflies in which the mutant gene is expressed specifically in a region known as the mushroom body. This area in the fly brain has been implicated in short-term olfactory learning⁶, as confirmed in gene-targeting experiments^{7,8}. The mushroom body is a central, multisensory processing structure, consisting of thousands of tightly packed neurons. These neurons receive information from nerve cells outside the mushroom body, and transmit signals to brain regions that are involved in controlling movements (Fig. 1). As mentioned above, the *shibire* mutant affects the transmission of signals, rather than the receipt of information. So, if this mutant is targeted to the mushroom body, it will affect only output (that is, neurotransmission) from the mushroom body, and not input. This makes it possible to identify the stages of learning that require neurotransmitter release from the mushroom body.

Dubnau et al.³ first shifted their mutant flies to the restrictive temperature, so impeding neurotransmission from the mushroom body, during memory acquisition (when they were learning to associate the odour with the shock) or during memory storage (the five minutes after training ended). Thirty minutes after training, the authors tested the flies for their recall of the association at the permissive temperature, when neurotransmission was normal. The flies showed normal retention of the memory. But when they were trained at the permissive temperature

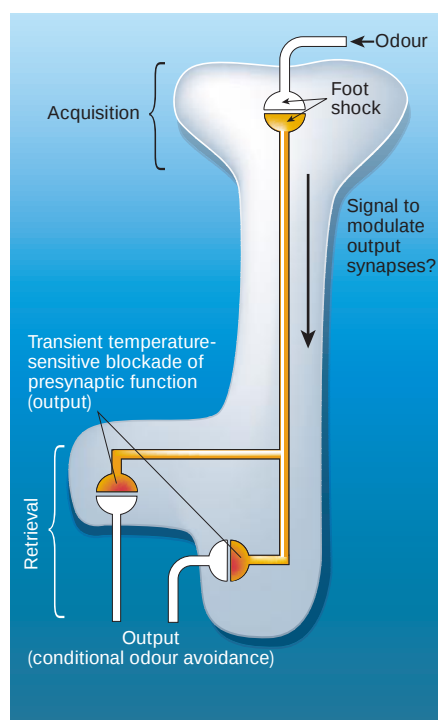


Figure 1 Teasing apart the different stages of learning in fruitflies. The mushroom body, pictured here, is part of the fruitfly brain, and is required for flies to learn to associate an odour with an electric shock in laboratory experiments. This process involves inputs to the mushroom body (neurons that carry the information of the odour and shock), and outputs (signals to the parts of the brain that coordinate motor processes, enabling the flies to respond to the odour). Dubnau et al.³ have used a temperature-sensitive mutant *shibire* gene to block output from mushroom-body neurons at different times — either while the flies are learning the odour–shock association, or while they are being tested for their recall of the association. They remember normally when output is blocked during learning, indicating that the output is not needed for memory storage. But the flies forget the association when the output is blocked during recall. This shows that the output is needed for memory retrieval, not acquisition.

and then tested at the restrictive temperature, memory retention was impaired.

So, memory acquisition and retrieval can be dissociated by interfering with neurotransmission in the mushroom body. The acquisition and early processing of memories are known to depend on a signalling cascade in the mushroom body that involves cyclic AMP⁹. This cascade seems to be driven by simultaneous input from olfactory interneurons and unknown modulatory neurons that are activated by the shock. This neural circuit, including input to the mushroom body and the mushroom body itself, might be a site for associative learning and memory storage. The sites of output from the mushroom body, however, are required for memory retrieval³ (at least for the early stages of memory tested here).

These results lead to several broad conclusions, as well as some questions. First, continuous neural activity at chemical synapses — often considered to be the basis of early memory phases — is not required in the mushroom body. So it seems unlikely that early memory formation requires any feedback from the mushroom body to input neurons, or to the mushroom body itself (that is, if electrical synapses are also not involved, a possibility not testable with the methods used here).

Second, the parts of mushroom-body neurons that are involved in memory acquisition and retrieval are clearly separate. But there must be some sort of signal that coordinates these processes. The signal probably includes the transport of molecules along mushroom-body neurons, and other, more rapid processes¹⁰. Dynamin is involved in protein transport as well as neurotransmission, so the memory-retrieval defects seen in the mutant flies might also reflect defects in coordination signals.

Third, a common feature of memory is its progression through different phases. The period tested by Dubnau *et al.*³ may represent a phase during which the formation of memories at mushroom bodies can be clearly dissociated from retrieval. But it is not known whether this separation also exists in later memory phases.

Finally, most memory traces in large brains are distributed over much greater distances than in fruitflies, and involve several areas of the brain¹¹. In the tiny fly brain, the memory trace of a rather simple, odour-guided behaviour is instead distributed between different sites in the same neuron, and these sites are specialized for the formation and retrieval of acquired information. Is this a design principle of a small brain, or of a simple type of memory, or both? The answers to these questions will help to show whether there is a general difference in how small and large brains cope with learning and remembering. ■

Randolf Menzel and Uli Müller are at the Institut

für Biologie–Neurobiologie, Freie Universität Berlin, Königin-Luise-Strasse 28–30, D-14195 Berlin, Germany.

e-mail: menzel@neurobiologie.fu-berlin.de

1. Ebbinghaus, H. *Memory: A Contribution to Experimental Psychology* (Dover, New York, 1885; reprinted in 1964).
2. Müller, G. E. & Pilzecker, A. Z. *Psychol.* **1**, 1–288 (1900).
3. Dubnau, J., Grady, L., Kitamoto, T. & Tully, T. *Nature* **411**, 476–480 (2001).

4. Tully, T. & Quinn, W. G. J. *Comp. Physiol. Psychol.* **156**, 263–277 (1985).
5. Dubnau, J. & Tully, T. *Annu. Rev. Neurosci.* **21**, 407–444 (1998).
6. Heisenberg, M. *Learning Memory* **5**, 1–10 (1998).
7. Connolly, J. B. *et al. Science* **275**, 2104–2107 (1996).
8. Zars, T., Wolf, R., Davis, R. & Heisenberg, M. *Learning Memory* **7**, 18–31 (2000).
9. Davis, R. L. *Neuron* **11**, 1–14 (1993).
10. Martin, K. C. *et al. Cell* **91**, 927–938 (1997).
11. Squire, L. R. *Memory and Brain* (Oxford Univ. Press, New York, 1987).

Condensed-matter physics

Why vortices matter

Peter Gammel

In a magnetic field, a superconductor is threaded by swirling whirlpools of electric current. Understanding these magnetic vortices is important because they control the flow of current through the superconductor.

The defining property of a superconductor is the ability to carry an electrical current without resistance when cooled below a transition temperature, typically a few degrees above absolute zero. But below this temperature, most of the superconductor's useful properties are governed by its response to external magnetic fields. Under certain conditions magnetic field lines can permeate some superconductors, leading to the creation of magnetic vortices. Vortices have non-superconducting cores, which carry the magnetic field lines, surrounded by swirling 'supercurrents'. The behaviour of these magnetic vortices determines the physical properties of superconductors, including the maximum electrical current the superconductor can support. The movement of vortices is particularly damaging because it creates electrical resistance, thereby destroying the superconducting state.

In 1986, a new class of high-temperature superconductor was discovered with unprecedented transition temperatures above 77 K. Magnetic vortices usually form stable and regular patterns, such as a triangular lattice, but in the high-temperature superconductors they appear in exotic and frequently less stable patterns. Two papers by Avraham *et al.*¹ and Bouquet *et al.*² (starting on page 448 of this issue) examine the response of vortices to changing temperatures and fields, and reveal some of their unusual behaviour in the high-temperature superconductors.

When a current flows in a high-temperature superconductor, it exerts a force on the magnetic vortices. If the vortices move in response to this force they dissipate energy and produce a non-zero resistance. But the vortices can also be fixed or 'pinned' to defects in the material, so they don't move. Practical applications of superconductors require zero resistance and therefore pinning. A solid hexagonal lattice was thought

to be the only stable vortex state until it was discovered that the vortex lattice can melt into a 'liquid' phase³. In the liquid phase it is more difficult to prevent vortex motion (and, hence, electrical resistance), which restricts applications of superconductivity to the vortex solid phases.

Like the melting of an ordinary solid, the melting of vortices is described as a first-order phase transition because it is accompanied by abrupt changes in physical

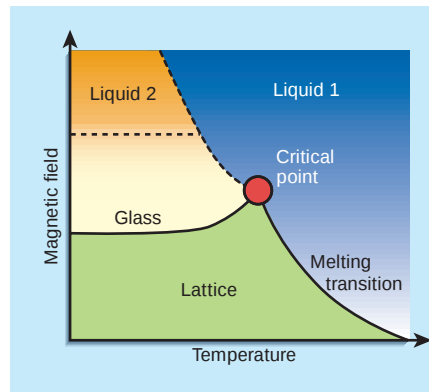


Figure 1 Phase diagram for magnetic vortices in a superconductor as a function of magnetic field strength and temperature. In the lattice and glass states, the superconductor will have zero electrical resistance. Motion of the vortices in the liquid states will generally lead to finite resistance. Taming the vortices is crucial for practical applications of superconductivity. Avraham *et al.*¹ show that the lattice–glass transition includes a region of 'inverse melting' just below the critical point. The critical point is where the solid, liquid and glass phases become indistinguishable. Bouquet *et al.*² have investigated the behaviour of vortices above the critical point, and identify a second transition between two distinct vortex liquids. These studies highlight the unusual properties of vortices in the high-temperature superconductors.

properties, such as magnetization. The properties of the vortices as a function of temperature and magnetic field can be summarized in a conventional phase diagram (Fig. 1). Detailed understanding of this phase diagram is essential for designing superconducting materials that have the properties needed for applications in microwave filters, magnetic field sensors and current-carrying cables.

The first-order melting transition from the vortex lattice to liquid 1 is well understood⁴ (solid line; Fig. 1). But the other parts of the phase diagram have remained a mystery, and this is where Avraham *et al.* and Bouquet *et al.* come in. Avraham *et al.*¹ study the first-order melting transition in the high-temperature superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (BSCCO), using microscopic Hall sensors to probe the changing magnetization. Previous studies of this transition in BSCCO were restricted to below the 'critical point' in Fig. 1, which occurs around a temperature of 40 K and a magnetic field of 25 millitesla (mT). The critical point is where the solid and liquid vortex phases become indistinguishable. Further exploration of the phase diagram above the critical point was hindered by the slow response of the lattice and glass (disordered solid) phases to changes in temperature and field.

Avraham *et al.* overcame this problem by applying a small alternating current (a.c.) field perpendicular to the existing magnetic field, to 'shake' the vortices into their new equilibrium state. They show that the first-order transition from lattice to glass is a continuation of the first-order vortex melting line. The slight upturn near the critical point in Fig. 1 means that the transition becomes 'inverse' — as the temperature increases, the disordered glass 'melts' into an ordered lattice. Inverse melting is a rare and counter-intuitive phenomenon⁵. For example, if melting of ice were inverse, an ice cube in a drink would actually grow as it cooled the surrounding liquid. Inverse melting cannot be a thermal process, because thermally driven transitions always proceed from less disorder at low temperature to increasing disorder at higher temperature.

Avraham *et al.*¹ suggest that the transition from the lattice to the glass phase is driven by pinning of the flux lines to impurities and disorder in the crystal. Such a transition would be almost temperature independent, but competition between thermal fluctuations and pinning disorder leads to the inverse melting near the critical point. This is a highly plausible explanation, but questions remain about the nature of the 'shaken' equilibrium compared to a real thermodynamic equilibrium. The unique properties of shaken equilibrium, for example in experiments with metal balls⁶, can lead to phase diagrams determined by the shaking

itself, rather than by a thermodynamic variable such as temperature.

Bouquet *et al.*² focus on the part of the phase diagram above the critical point, which corresponds to high magnetic fields. They measure both heat capacity and magnetization for defect-free crystals of the high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_7$ (YBCO). In YBCO, superconductivity extends to much higher temperatures and fields than in BSCCO, so that the critical point is at a temperature of 75 K and a magnetic field of 11 T. The purity of the crystal used by Bouquet *et al.* is shown by the nature of the first-order melting transition below the critical point. Below 5 T, the vortex melting is sharp (first order), and the heat capacity and magnetization are precisely related. The melting transition remains first order up to 11 T, although between 5 T and 11 T thermal fluctuations and disorder start to compete. At fields above the critical point, up to 26 T, Bouquet *et al.* identify a second-order transition between two distinct types of vortex liquid (liquid 1 and liquid 2 in Fig. 1). A second-order transition is less sharp, such that disorder and other properties change continuously across the transition.

To understand the distinction between these two vortex liquids, Bouquet *et al.* measured their heat capacity. Their data agree with a theory⁷ in which the magnetic field lines in the low-temperature vortex liquid are thought of as having 'line tension', as if they were rubber bands. In this picture, liquid 1 has line tension, whereas liquid 2 does not. This means that the flux lines in liquid 1 can be pulled, elongated and twisted without losing energy.

The work of Avraham *et al.* and Bouquet *et al.* gives us a more comprehensive view of the vortex phase diagram. Exactly how the pieces of the phase diagram they explore join together at the critical point is still uncertain and is likely to depend on the detailed structural disorder of the materials⁸. Away from the critical point, and armed with the insight provided by these experiments, we can hope to control the vortices in both the glass and liquid phases. This will eventually allow superconducting devices to operate at much higher temperatures and magnetic fields than before. ■

Peter Gammel is in the Electronic Devices Research Laboratory, Agere Systems, Murray Hill, New Jersey 07974, USA.

e-mail: plg@agere.com

1. Avraham, N. *et al.* *Nature* **411**, 451–454 (2001).
2. Bouquet, F. *et al.* *Nature* **411**, 448–451 (2001).
3. Bishop, D. J., Gammel, P. L., Huse, D. A. & Murray, C. A. *Science* **255**, 165–172 (1992).
4. Zeldov, E. *et al.* *Nature* **375**, 373–376 (1995).
5. Greer, A. L. *Nature* **404**, 134–135 (2000).
6. Umbanhowar, P. B., Melo, F. & Swinney, H. L. *Nature* **382**, 793–796 (1996).
7. Tesanovic, Z. *Phys. Rev. B* **59**, 6449–6474 (1999).
8. Safar, H., Gammel, P. L., Huse, D. A. & Bishop, D. J. *Phys. Rev. Lett.* **70**, 3800–3803 (1993).

Daedalus

Gas is for burning

Combustion, says Daedalus, is the most fundamental reaction for disposing of human rubbish. And yet modern civilization uses it very badly. A generation or two ago, the universal coal or coke fire automatically got rid of old newspapers and scrap food; today it would easily remove waste plastics as well. Yet these all go to fill our dustbins. The contents are then incinerated, which is highly unpopular, and finally as landfill they evolve combustible methane gas, a useful product we cannot collect and use safely. There must be a better way.

The better way which DREADCO engineers are studying is the 'Burnall', a device designed to burn to gas, waste and anything else that comes its way. It was inspired by Daedalus' memory of the gas poker, which one placed in the coke to start a fire. The idea is that gas is burned in a normal grate; immediately below this is a tray on which is placed the shredded paper waste, food waste or other waste. Heat radiated downwards gasifies these unwanted items, and their vapour burns in the gas stream; their carbon goes more slowly to one of the oxides. All the skill of domestic engineering will be needed to make this consumer durable effective and acceptable; but the thing seems feasible. Many of the problems have already been solved on a large scale, in hospital incinerators for example.

A steady stream of unfortunate people gas themselves by cooling a gas flame in a water-coil (thus producing deadly carbon monoxide); so the Burnall will be topped by a mesh of appropriate catalyst to convert all such nasties to harmless carbon dioxide and nitrogen. The output stream of gas, maybe at 400 °C, will traverse water coils to transfer its heat to a central heating unit. If it is worth it, a second condensation stage will capture the latent heat of condensation of steam, and pass the water to a drain.

The DREADCO Burnall will effortlessly solve many of the problems of modern living. All the smelly troublesome wastes will now escape the dustbin, and will go up the flue to reduce the heating bill. All the tedious paper and card, plastic wrappings, soiled infants' disposable nappies, and so on, will be reclaimed for their heat and saved from the binman. Sadly, bottle-tops, buttons and other objects containing PVC and related plastics will liberate unwanted chlorine. A controllable feed of lime, converting this into calcium chloride in the white and harmless ash, may be necessary.

David Jones

Obituary

John Frank (Jack) Allen (1908–2001)

In late 1937, two young physicists at the Royal Society Mond Laboratory in Cambridge, UK, found that liquid helium could flow through very small capillaries with essentially zero viscosity below a temperature of 2.17 kelvin. While they were preparing a note for publication, they heard of another paper, just submitted by P. L. Kapitza in Moscow, reporting similar results. Both papers appeared side by side on 8 January 1938 in *Nature*. Kapitza and the young physicists Jack Allen and A. D. Misener had discovered the mysterious phenomenon of superfluidity in liquid ^4He . It was the start of a golden period in low-temperature physics. Allen, the last survivor of this heroic time, died of a stroke on 22 April 2001, aged 92.

Most of Allen's greatest discoveries were made at the outset of his career (in 1938 he was essentially a postdoc, working without a supervisor, which suited him fine). Between 1937 and 1939 he and his associates at Cambridge produced a stream of papers on superfluidity in liquid ^4He , this output coming to an abrupt end with the start of the Second World War. In 1946, with normal life resuming, Allen organized the first international low-temperature meeting at the University of Cambridge, UK. This was the birth of the major triennial event in the low-temperature physics community; the twenty-third meeting will be held in Hiroshima next year.

In 1947, Allen was appointed professor of natural philosophy and head of physics at the University of St Andrews in Scotland, and two years later he became a Fellow of the Royal Society. But although he stayed active in the world of low-temperature physics, revitalizing research at St Andrews and helping to run many conferences and workshops, his own research was never again central. He retired from St Andrews in 1978.

Jack Allen was born in Winnipeg, Canada, in 1908 (the year ^4He was first liquefied in Leiden, in the Netherlands). He received his BA from the University of Manitoba, where his father was head of the physics department. In 1929, he went to study at the University of Toronto where James F. McLennan had built up a strong physics department. In 1923 this had been only the second laboratory in the world to liquefy helium, and when Allen arrived he immersed himself in the new world of cryogenics. He received his PhD in 1933, having already written ten papers on superconductivity. Late in 1935 he went



Co-discoverer of superfluidity

to Cambridge to work with Kapitza, only to find that Kapitza was being detained in Moscow, and was in the process of setting up what was to become the famous Institute of Physical Problems. In Kapitza's absence, Allen effectively took over the low-temperature work, although officially the director was J. D. Cockroft.

In 1935, Don Misener, a graduate student at Toronto, had carried out the first experimental study of the viscosity of liquid ^4He . By then it was known that liquid helium underwent some sort of phase transition at 2.17 K, as there were abrupt changes in various thermodynamic properties. Misener's work suggested that the viscosity decreased substantially when the liquid passed through this transition. Misener joined Allen at Cambridge in 1937 to do his PhD, and the two set out to study the phenomenon by examining flow in thin capillaries.

Misener's 1935 experimental work had also attracted the notice of Kapitza in Moscow. Both groups reported their independent discovery of superfluid flow in 1938, with Kapitza being the first to coin the term 'superfluid'. It is puzzling and unfortunate that when Kapitza finally received a well-deserved Nobel prize in physics in 1978, the citation concerning superfluidity made no reference to the work of Allen and Misener.

Allen quickly found other dramatic manifestations of superfluidity, all of which involved the counterflow of the normal and superfluid components or the 'clamping' of the normal component with

fine powder. But after 1945 the Moscow group under Kapitza (helped by L. D. Landau, who developed a complete theory of the two-fluid behaviour of superfluid helium in 1941) dominated further research on quantum liquids. The study of superfluid ^4He increasingly involved microscopic theories and new experimental probes such as neutron scattering, none of which interested Allen.

Allen was the last of a generation of independent-minded classical physicists who delighted in explaining the visible world. He prized his own ability, and that of glass-blowers and technical people, to build experimental apparatus. Allen was as proud of his invention of the O-ring vacuum seal as anything else he did. It should be no surprise that his greatest work on superfluidity in liquid ^4He involved phenomena that could be seen. Indeed, it is most fitting that Allen discovered the famous 'fountain effect' in 1938 with the help of a pocket flashlight.

Over a ten-year period Allen made a movie of the various two-fluid phenomena exhibited by liquid ^4He . The photography of these effects was a real challenge, because liquid ^4He is essentially transparent. This unique colour movie (the fifth edition was completed in 1982) is one of Allen's great legacies to physics.

Allen had a commanding presence and a dry sense of humour. He strongly identified with the physics of an earlier day, and I imagine that he would have enjoyed talking with classical physicists such as Lord Rayleigh, Michael Faraday and Daniel Bernoulli more than with Werner Heisenberg and Erwin Schrödinger. Superfluidity is a dramatic visible manifestation of quantum mechanics, being the result of Bose–Einstein condensation in which a macroscopic number of ^4He atoms occupy the same, single-particle quantum state. It is paradoxical that the phenomenon was first observed by Allen: a great physicist who wasn't much interested in atoms.

Walking the old stone streets of St Andrews, one quickly notices the elegant metal historical plaques on many buildings, commemorating famous people who have been associated with this historic town and its university over the centuries. Almost every plaque connects its subject to physics — not surprising, because Jack Allen was the motivating force behind the sign committee. I very much hope that the town sees fit to so honour Allen himself.

Allan Griffin

Allan Griffin is in the Department of Physics, University of Toronto, Toronto, Ontario, M5S 1A7, Canada.
e-mail: griffin@physics.utoronto.ca

Converting currencies in the Old World

Simple arithmetic underpinned trading throughout the Near East during the Bronze Age.

Raw materials recovered from archaeological excavations in the Indus Valley, the Persian Gulf, Mesopotamia, Egypt and the eastern Mediterranean reflect the existence of long-distance trading during the Bronze Age, which united these regions into networks of commercial exchange. As each region relied on a different set of weights for trading, a straightforward conversion system must have been in operation. Here we describe a simple and universal conversion system that could have provided an economic key to the trade networks of the Old World between 2500 and 1000 BC.

Old World trading networks brought raw resources such as lapis lazuli, copper, tin, silver and gold, as well as exotic materials and artefacts such as cylinder seals, decorative pottery and carved stone vessels from far afield. Establishing equivalence between weight systems that were peculiar to different regions would have created comparable units of value as well as mutual confidence in buying and selling goods, which would have been essential for dealing with universally precious materials such as gold and silver.

Many different weight systems were in operation in the Bronze Age (see Fig. 1 for an example), all of which were more or less amenable to a 'standard' system of conversion. A Mesopotamian text¹ from about 1800 BC mentions the conversion of a weight from the Dilmun standard (used on the islands and northeastern shores of the Arabian peninsula) to the Old Babylonian Ur standard: "... they had given to us 5,532 $\frac{2}{3}$ minas of copper according to the standard of Dilmun. Its weight is in total 611 talents 6 $\frac{1}{2}$ minas according to the standard of Ur."

Different approaches have been used to study ancient weight systems²⁻⁹ — for example, attempts have been made to trace a linear development into our prevailing metrological systems². We base our investigation on the use of talents, shekels and minas as units of weight, although their actual weights vary in different systems (see Table 1 of supplementary information). For example, a Syrian talent is 28.2 kg, which is subdivided into 60 minas of 470 g each; a mina is subdivided into 50 shekels of 9.4 g. Thus, in the Ugarit system, 3,000 shekels of 9.4 g ($3,000 \times 9.4 \text{ g} = 28,200 \text{ g}$) would be the same as a Syrian talent of 28.2 kg.

This Ugaritic shekel of 9.4 g can be inter-related with all other weight systems in the Near East and the Mediterranean. It is identical to the Egyptian kdt of 9.4 g and is equal to three-quarters of the Egyptian gold



Figure 1 Stone weights in the shape of ducks from the archaeological site of Yorgha Tepe, ancient Nuzi, Iraq. Nuzi was a provincial town of the Hurrian Kingdom (1600–1400 BC). The weights represent 20 shekels (smallest), 1 mina and 2 minas (largest), and are marked with incisions to denote their value. (Photo: Semitic Museum, Harvard University.)

dbn of 12.83 g. Furthermore, four Ugaritic shekels ($4 \times 9.4 \text{ g} = 37.6 \text{ g}$) is roughly equivalent to three Egyptian gold dbn ($3 \times 12.83 \text{ g} = 38.49 \text{ g}$). We doubt that the weight systems in the second millennium BC were accurate enough to detect an error of less than 1 g in 38 g, or 2.6%.

These equivalences allowed the use of weights in different combinations, generally multiples of two or five. For example, two weights each with the value of two Ugaritic shekels were equivalent to two weights of the value of one and two Egyptian dbn. Thus, two Ugaritic shekels ($9.4 \text{ g} \times 2 = 18.8 \text{ g}$) plus two more Ugaritic shekels comes to 37.6 g, and two weights with the value of one (12.8 g) and two Egyptian gold dbn (25.6 g) comes to 38.4 g. In another example, 9.4 g of the Ugaritic shekel equals four-fifths of the Hittite shekel of 11.75 g ($5 \times 9.4 \text{ g} = 47 \text{ g}$ and $4 \times 11.75 \text{ g} = 47 \text{ g}$), so five Ugaritic shekels equals two weights each with the value of two Hittite shekels.

This conversion system enables a hypothetical 1,370-g 'ingot' of lapis lazuli to be defined in terms of the different shekel weights of each region. Thus, 1,370 g would be the equivalent of 100 13.68-g Dilmun shekels, 160 8.55-g Mesopotamian shekels, 175 7.83-g Eblaite–Carcemish shekels, and so on, across an east–west route to the Mediterranean (see Table 2 of supplementary information, which also uses a tin ingot as an example).

Our analysis of the diverse weight systems that were in use from the Indus Valley to the Aegean from the middle of the third

millennium to the end of the second millennium BC complements earlier studies⁶⁻⁸ and reveals that the conversion systems in operation were elegant in their simplicity, although each weight system showed a degree of variance from our proposed standards. The integration of different weight systems was crucial in developing the scale and nature of commercial exchange in the Near East and would have facilitated the emergence of the Ancient World System¹⁰.

Alfredo Mederos*†, C. C. Lamberg-Karlovsky†

**Departamento de Prehistoria, Universidad Complutense, Ciudad Universitaria, 28040 Madrid, Spain*

*†Department of Anthropology, Peabody Museum, Harvard University, 11 Divinity Avenue, Cambridge, Massachusetts 02138, USA
e-mail: karlovsk@fas.harvard.edu*

1. Roaf, M. *Iraq* **44**, 137–141 (1982).
2. McDonald, D. M. *The Origins of Metrology* (McDonald Institute for Archaeological Research, Cambridge, 1992).
3. Bibby, T. G. *Kumli* **1970**, 345–353 (1970).
4. Powell, M. A. in *Studies in Honour of T. B. Jones* (eds Powell, M. A. & Sack, R. H.) 71–201 (Kevelaer, 1979).
5. Edzard, D. O. (ed.) *Reallexikon der Assyriologie und Vorderasiatischen Archäologie* Vol. 7, 457–517 (de Gruyter, Berlin, 1990).
6. Parise, N. F. *Dialoghi di Archeologia* **3**, 155–159 (1981).
7. Zaccagnini, C. in *Traffici Micenei nel Mediterraneo* (eds Marazzi, M. et al.) 303–314 (Istituto per la Storia e l'Archeologia della Magna Grecia, Taranto, 1986).
8. Zaccagnini, C. *II Congresso Internazionale di Studi Fenici e Punici* 343–347 (CNR, Roma, 1991).
9. Courtois, J. C. *Res. Orientales* **2**, 119–127 (1990).
10. Frank, A. G. *Curr. Anthropol.* **34**, 383–429 (1993).

Supplementary information is available on Nature's website at <http://www.nature.com> or as paper copy from the London editorial office of Nature.

Plant pathogens

Mitochondrial control of fungal hybrid virulence

Hybrid species of fungal pathogens that infect wild and cultivated plants are emerging with new virulence and host ranges, posing a threat to agriculture and forestry¹. Here we show that the virulence of hybrid species of the basidiomycete fungus *Heterobasidion annosum* (Fr) Bref, a causal agent of root and butt rot in conifers and one of the most economically important forest pathogens, is controlled by their mitochondrial genome. Our results indicate that cooperation between organelles that contain genetic information may influence the phenotype of hybrid phytopathogens.

Among fungal plant pathogens, species hybrids have been discovered in *Melampsora*^{2,3}, *Phytophthora*^{4–6}, *Ophiostoma*⁷ and *Heterobasidion*⁸. In North America there are two intersterile groups of *H. annosum* that correspond to species with overlapping morphological characters and different ecological niches: the 'S' group infects *Abies* spp. and *Tsuga* spp.⁹, whereas the 'P' group has *Pinus* spp. as its main hosts⁹. So far, only one natural S–P hybrid has been found⁶ — this isolate is less virulent than either of its progenitors to their 'adapted' host plant species but not to a universal host¹⁰.

We artificially created four S–P hybrid heterokaryons (fungal mycelia that contain dissimilar nuclei), one S–S heterokaryon and one P–P heterokaryon from four North American homokaryotic *H. annosum* isolates (containing one nuclear type each; provided by T. E. Chase) and determined the severity of their effects on pine seedlings (Table 1).

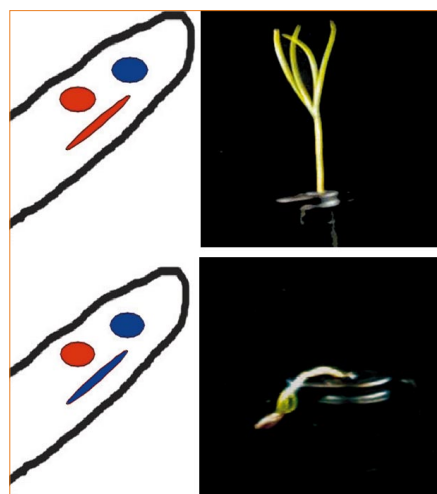


Figure 1 Virulence of S–P hybrids of the plant pathogen *H. annosum* to pine seedlings. Virulence of hybrids is controlled by the mitochondrion (represented by ellipses) rather than the nucleus (circles). Blue denotes genomic inheritance from the P group, red from the S group. Hybrids with mitochondria of S origin do not kill the pine seedling (top), in contrast to hybrid mycelia with mitochondria of P origin (bottom).

Table 1 Virulence of fungal homokaryons and heterokaryons

Isolate	Genotype	Nucleus type	Mitochondrial type	Mortality (%)
TC 32-1	A	P	P	90
TC 39-7	B	P	P	90
TC 122-11	C	S	S	31
TC 122-12	D	S	S	34
AO 8	A+D	S+P	S	55
AO 9	B+D	S+P	P	86
AO 10	A+B	P+P	P	93
AO 11	C+D	S+S	S	28
AO 12	B+C	S+P	S	28
AO 13	A+C	S+P	P	83

Genotypes of the homokaryons and their combinations in heterokaryons are represented by the letters A–D. Nuclear types were confirmed by mating tests and by microscopic examination for clamp connections, which indicate heterokaryons. Mitochondrial types were determined from the intron size in the gene that encodes large-subunit ribosomal RNA (see text). Virulence is expressed as mortality rate among 29 pine seedlings for each isolate 20 days after inoculation with 1.7 ml of mycelium suspension (0.5 mg dried mycelia per ml of medium).

We analysed the virulence of the homokaryons and heterokaryons *in vitro* by visually scoring dead or live pine seedlings daily for 20 days¹¹. We observed no significant difference in virulence between homokaryons and heterokaryons of either the S or P intersterility groups ($P < 0.05$, χ^2 test). This is consistent with earlier results suggesting that heterokaryosis of *H. annosum* is not required for virulence¹².

We scored S–P hybrids as having either low or high virulence to pine, which is characteristic of the S and P groups, respectively (Table 1). Isolates AO 9 and AO 13 did not differ significantly in virulence from P isolates, but were significantly different from isolates AO 8 and AO 12, and from S isolates ($P < 0.05$, χ^2). Neither AO 8 nor AO 12 differed significantly from S isolates, but they were different from P isolates ($P < 0.05$, χ^2). If the virulence of a *H. annosum* hybrid were solely controlled by nuclear virulence genes or by avirulence genes, then all S–P hybrids should show either the high or low virulence that characterizes the respective intersterility groups.

Homokaryotic *H. annosum* S and P isolates differed in the size of an intron in the gene that encodes the large-subunit RNA of mitochondrial ribosomes, as revealed by amplification with the ML5–ML6 primer pair⁸ (Table 1). The mitochondria of the hybrid isolates AO 8 and AO 12 were from the S homokaryon, whereas those of AO 9 and AO 13 were of P origin (Table 1). None of the S–P hybrids possessed both types of mitochondrion, showing that mitochondrial transmission was uniparental.

There was a perfect correlation between the mitochondrial type acquired by the hybrids and their virulence (Table 1; Fig. 1). Hybrids with mitochondria of P origin had similar virulence to P homokaryotic progenitors, whereas the virulence of hybrids with S mitochondria resembled that of S homokaryons. To our knowledge, this is the first example of mitochondrial control of virulence in species-hybrid plant pathogens, although mitochondria are thought to be important in host specialization for *Phy-*

tophthora nicotianae × *P. cactorum* hybrids⁵.

What is the source of this correlation between mitochondrial origin and the virulence of pathogenic fungal hybrids? It may be that the mitochondrial genome itself encodes factors that determine the virulence of hybrids, or that interplay between the organism's mitochondrial and nuclear genomes gives rise to a mitochondrially determined phenotype in hybrids.

Åke Olson, Jan Stenlid

Department of Forest Mycology and Pathology,
Swedish University of Agricultural Sciences,
SE-750 07 Uppsala, Sweden
e-mail: ake.olson@mykopat.slu.se

- Brasier, C. *Nature* **405**, 134–135 (2000).
- Spier, A. G. & Hopcroft, D. H. *Mycol. Res.* **98**, 889–903 (1994).
- Newcombe, G., Stirling, B., McDonald, S. & Bradshaw, H. D. Jr *Mycol. Res.* **104**, 261–274 (2000).
- Éresk, T., English, J. T. & Schoelz, J. E. *Phytopathology* **85**, 1343–1347 (1995).
- Man in't Veld, W. A. et al. *Phytopathology* **88**, 922–929 (1998).
- Brasier, C. M., Cooke, D. E. L. & Duncan, J. M. *Proc. Natl. Acad. Sci. USA* **96**, 5878–5883 (1999).
- Brasier, C. M., Kirk, S. A., Pipe, N. D. & Buck, K. W. *Mycol. Res.* **102**, 45–57 (1998).
- Garbelotto, M., Ratcliff, A., Bruns, T. D., Cobb, F. W. & Otrosina, W. J. *Phytopathology* **86**, 543–551 (1996).
- Korhonen, K. & Stenlid, J. in *Heterobasidion annosum: Biology, Ecology, Impact and Control* (eds Woodward, S., Stenlid, J., Karjalainen, R. & Hüttermann, A.) 43–70 (CAB International, Wallingford, 1998).
- Garbelotto, M., Popenuck, T., Ratcliff, A., Cobb, F. W. & Bruns, T. D. *Phytopathology* **86** (suppl.), 28–29 (1996).
- Sharma, P., Börja, D., Stougaard, P. & Lönneborg, A. *Physiol. Mol. Plant Pathol.* **43**, 57–67 (1993).
- Garbelotto, M. et al. *Mycologia* **89**, 92–102 (1997).

addendum

A fern that hyperaccumulates arsenic

L. Q. Ma, K. M. Komar, Cong Tu, Weihua Zhang, Yong Cai, E. D. Kennelley
Nature **409**, 579 (2001)

It has been pointed out that other plants have been reported previously to take up and tolerate arsenic from soils^{1–5}, but we would like to stress that, unlike brake fern (*Pteris vittata*), none of these plants has been shown to hyperaccumulate arsenic — that is, the concentration of arsenic in their aboveground biomass was less than that in the soil and in the plant roots.

- Porter, E. K. & Peterson, P. J. *Sci. Tot. Env.* **4**, 365–371 (1975).
- De Koe, T. *Sci. Tot. Env.* **145**, 103–109 (1994).
- Bech, J. et al. *Sci. Tot. Env.* **203**, 83–91 (1997).
- Warren, H. V., Delavault, R. E. & Barakso, J. *Can. Min. Metal. Bull.* **61**, 860–867 (1968).
- Meharg, A. A., Naylor, J. & Macnair, M. R. *J. Env. Qual.* **23**, 234–238 (1994).

Switch-based mechanism of kinesin motors

Masahide Kikkawa^{*†}, Elena P. Sablin^{†‡}, Yasushi Okada^{*}, Hiroaki Yajima^{*}, Robert J. Fletterick[‡] & Nobutaka Hirokawa^{*}

^{*} Department of Cell Biology and Anatomy, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan

[‡] Department of Biochemistry/Biophysics, University of California, San Francisco, California 94143-0448, USA

[†] These authors contributed equally to this work

Kinesin motors are specialized enzymes that use hydrolysis of ATP to generate force and movement along their cellular tracks, the microtubules. Although numerous biochemical and biophysical studies have accumulated much data that link microtubule-assisted ATP hydrolysis to kinesin motion, the structural view of kinesin movement remains unclear. This study of the monomeric kinesin motor KIF1A combines X-ray crystallography and cryo-electron microscopy, and allows analysis of force-generating conformational changes at atomic resolution. The motor is revealed in its two functionally critical states—complexed with ADP and with a non-hydrolysable analogue of ATP. The conformational change observed between the ADP-bound and the ATP-like structures of the KIF1A catalytic core is modular, extends to all kinesins and is similar to the conformational change used by myosin motors and G proteins. Docking of the ADP-bound and ATP-like crystallographic models of KIF1A into the corresponding cryo-electron microscopy maps suggests a rationale for the plus-end directional bias associated with the kinesin catalytic core.

Although most kinesin motors function as multisubunit assemblies, KIF1A is composed of a single polypeptide chain (Fig. 1a) and is a monomer under *in vivo* conditions^{1,2}. KIF1A is the processively moving monomeric kinesin taking multiple steps before detaching from microtubules³. The catalytic core of KIF1A (comprising less than 25% of the motor and sharing 43% amino-acid sequence identity with the motor domain of conventional kinesin) is processive when attached to the neck linker of either its own or conventional kinesin³. Here, we have determined the atomic structure of the KIF1A catalytic core (with part of the neck linker of conventional kinesin; Fig. 1a) in its two functionally important states: with ADP and with the non-hydrolysable ATP analogue β , γ -methyleneadenosine 5-triphosphate (AMPPCP) bound in its active site. Both ADP-bound and ATP-like (AMPPCP-bound) structures were solved using molecular replacement methods and refined to 2.2 and 2.0 Å, respectively.

The KIF1A catalytic core in both states has the same overall architecture, which resembles that of the other kinesins^{4–10}. Despite the overall structural similarity of the cores, four surface loops in KIF1A (L2, L3, L10 and L12) differ from the corresponding loops of conventional kinesin⁶. L3 is remarkable in that only the monomeric kinesins have a conserved insertion in this loop. As L3 forms a part of the nucleotide-binding pocket, it might modulate nucleotide binding and release during the enzymatic cycle of monomeric kinesins. L12, which constitutes a principal component of the microtubule-binding interface of the kinesin core^{11–15}, in the case of KIF1A, has an insertion of 12 residues of which 6 are lysines. This positively charged 'K-loop' is uniquely conserved among monomeric kinesins and is essential for their processivity^{15,16}. Although the carboxy-terminal part of L12 is universally conserved (residues 303–307 in KIF1A) and well ordered in both of the KIF1A structures, the K-loop insertion is flexible and not visible in our electron density maps (Fig. 1b).

Conformational switching between ADP and ATP-like states

To identify regions in the catalytic core that change their conformation between different nucleotide-bound states—and are therefore potentially crucial for motor function—the ADP-bound and the ATP-like KIF1A structures have been superimposed using the C α atoms of the conserved nucleotide-binding loop, the P-loop (residues 97–104), which remains unchanged throughout the

nucleotide hydrolysis cycle¹⁷. Although the two structures look very similar overall, two regions in the KIF1A catalytic core show marked differences between the ADP and ATP-like states (Fig. 1b). The first consists of a single loop, L9 (residues 202–218; r.m.s. deviation of 6.2 Å for 17 C α atoms). The second is composed of five consecutive structural elements that form a cluster in three-dimensional space. These are loop L11, helix α 4, loop L12, helix α 5 and loop L13, called the 'switch II cluster' (residues 248–324, r.m.s. deviation of 5.1 Å for the 48 C α atoms that are visible; for comparison, r.m.s. deviation for the remaining 255 C α atoms of the core is 0.9 Å).

These two regions include the switch I and switch II elements (with consensus sequence motifs NXXSSRS (amino acids 211–218) and DLAGXE (residues 248–253), respectively; where 'X' indicates any amino acid) named on the basis of their similarity to the structural counterparts in G proteins^{5,18}, the well-known molecular switches that regulate protein/protein interactions in many biological systems. It was thought that, similar to G proteins, the switch regions of kinesins and myosins might sense the γ -phosphate in bound nucleotide and trigger nucleotide-dependent conformational change in the motors^{5,18}. For myosins, this prediction was verified by comparing structures of myosins in different nucleotide-bound states^{19–22}. For kinesins, this hypothesis is confirmed in the present analysis, which clearly reveals the two switch regions adopting conformations that depend on the nature of the bound nucleotide (Fig. 1b).

The switch I region of KIF1A forms a helix flanked by two short loops in the ADP state. In the ATP-like state, the same region forms a short β -hairpin structure (Fig. 1b). Similar to G proteins and myosins, the restructuring of the polypeptide backbone in this area brings the switch I region of the kinesin core closer to the nucleotide in the ATP-like state compared with the ADP-bound structure. The conformational change of switch I in KIF1A might be triggered by a special conserved serine (S215, similar to the conserved serine of myosins¹⁹ and threonine of G proteins²³), which in the ATP-like state can interact with the γ -phosphate of the nucleotide (Fig. 1c). Although binding contacts with the γ -phosphate define the position of the C-terminal base of the switch I loop (residues 214–217), the remaining part of this loop could show significant flexibility in the ATP-like state. Structural comparison of various kinesins with bound ADP also show considerable variation

in the conformation of the switch I region²⁴. Similar malleability of the switch I loop in G proteins accommodates their interactions with macromolecular partners, which structure the switch I region^{17,25,26}. The exact structure of the kinesin switch I loop might be defined by binding of the motor to the microtubule.

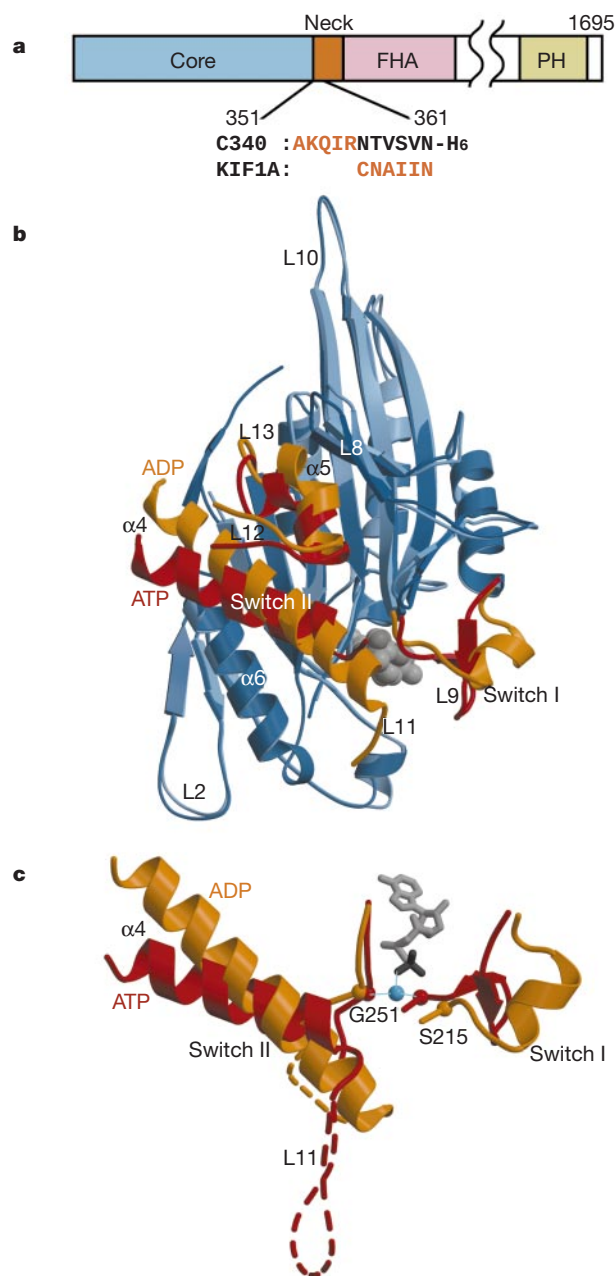


Figure 1 Structural analysis of KIF1A in the ADP and ATP-like states. **a**, Domain organization of the KIF1A polypeptide chain. The N-terminal catalytic core (blue) is adjacent to the class-specific neck region (orange)³⁴. The following ‘fork-head-associated’ (FHA, pink) domain is specific for monomeric kinesins and might bind to the microtubule. The C-terminal pleckstrin homology (PH, yellow) domain is involved in cargo binding. Orange and black letters indicate the N-terminal part of the KIF1A neck linker and the attached part of the conventional kinesin KIF5C neck linker, respectively. **b**, Superposition of ADP (light blue) and the ATP-like (dark blue) KIF1A structures. Regions of the core that undergo large conformational changes are shown in yellow and red for ADP and the ATP-like states, respectively. **c**, Nucleotide-driven conformational changes in the KIF1A switch regions. The disordered part of L11 is shown as yellow and red dashed lines; the γ -phosphate is in dark grey. Switch I serine (S215) and switch II glycine (G251) are shown as yellow and red spheres; bridging water molecule is indicated as a blue sphere.

The switch II region of KIF1A, which constitutes a part of the switch II cluster and includes loop L11 and helix $\alpha 4$, also undergoes a nucleotide-dependent conformational change (Fig. 1b) similar to that observed for the switch II region of G proteins and myosins. As the structural comparison of the two KIF1A structures shows, a conserved glycine at the base of the switch II loop L11 (G251, equivalent to the conserved glycine of myosins¹⁹ and G proteins²³) probably triggers the conformational change in the kinesin switch II region. In the ATP-like structure of KIF1A, the amide group of this glycine can interact with the γ -phosphate of the nucleotide (Fig. 1c). The nucleotide-induced movement of the base of L11 would rearrange the structure and change the position of the entire switch II cluster, as it is composed of closely interacting, consecutive structural elements.

Although in our crystallographic models of KIF1A most of L11 is disordered and not visible (Fig. 1b, c), this loop is a principal component of the kinesin microtubule-binding interface^{12,15} and is structured in the microtubule-bound state of KIF1A. Analogously, the inherently unstable and malleable switch II region of G proteins²³ is stabilized and structured when G proteins form complexes with their macromolecular partners^{25,26}. L11 of KIF1A undergoes a change in its length and a related change in its secondary structure (Fig. 1c). These seem to be crucial for relaying the signal from the nucleotide-binding site. In the ATP-like state, L11 is lengthened compared with the ADP-bound structure because the first two amino-terminal turns of the following switch II helix $\alpha 4$ are unwound (Fig. 1b, c). In addition, the switch II helix in the ATP-like state undergoes translation towards the nucleotide (~ 4 Å) and rotational movement ($\sim 20^\circ$) away from the central β -sheet and towards the C-terminal helix $\alpha 6$. As a result, three interlocked consecutive structural elements (L12/ $\alpha 5$ /L13, the remaining part of the switch II cluster) all follow the same direction (Fig. 1b). The latter movement is facilitated by the presence of conserved flexible glycines (G320 and G321) in L13. In the ADP state the entire switch II cluster slides over the central β -sheet away from the nucleotide as the switch II loop L11 retracts. The retraction follows from the formation of two additional turns of the switch II helix. Similar nucleotide-induced coil/helix transitions have been reported for the switch II region of G proteins^{17,27}.

Because the nucleotide-dependent conformational changes in the KIF1A catalytic core involve the switch regions (preserved in kinesins, myosins and G proteins^{5,18}) these movements probably predict a general mechanism of nucleotide-driven conformational switching for all kinesins. Although the presence of the γ -phosphate in the bound nucleotide coordinates the movements of the switch regions in kinesins, these regions can move in the absence of the γ -phosphate²⁸. When all structures of kinesin catalytic cores^{4–10} are superimposed they form two distinct categories. The switch II cluster of most of the kinesin structures overlaps precisely with the corresponding cluster in the KIF1A ADP-complexed structure (this is not unexpected, as all of them were solved in complex with Mg-ADP). However, the switch II cluster in rat conventional kinesin^{6,7}, with ADP bound, superimposes with this region in the ATP-like state of KIF1A. Under different crystallization conditions, the ATP-like conformation of the switch II cluster was observed in a distinct KIF1A/Mg-ADP crystal structure (see Methods). These results can be explained by a low-energy barrier between two conformations of the switch II cluster, which might exist in dynamic equilibrium in the absence of the γ -phosphate and the microtubules²⁸. This hypothesis is consistent with the idea that myosin transitions between the ADP and ATP-like states occur without an appreciable activation energy²². Both the myosin and kinesin motors, therefore, might switch easily between two states in response to external forces (crystallization conditions or crystal-packing contacts). The choice of conformation in the working cycles of the motors is probably decided by binding contacts with both the nucleotide and the track polymer.

Conformational transitions of the mechanical elements

The conformational changes in the switch regions of some G proteins trigger massive domain rearrangements^{29,30} that are related to specific functions performed by a particular class of G protein. In myosin motors, the nucleotide-driven rearrangements of the switch regions result in the angular motion of the lever-arm domain, which powers myosin motility^{20–23}. We have found that the nucleotide-dependent movements of the kinesin switch II cluster result in dramatic rearrangements of the kinesin neck linker, which is located at the C terminus of the catalytic core (Fig. 1a). As kinesin neck domains have been implicated in force-generating process and movement of different kinesin motors^{9,31–36}, we propose that the switch-controlled domain movements in kinesins evolved to power their motility. In the KIF1A structures, the ATP-like conformation of the switch II cluster allows the neck linker to be docked on the catalytic core (Figs 1b and 2a). A different position of the same cluster in the ADP state occludes such docking, pushing the neck linker away from the core. As a result, the neck linker is docked on the core in the ATP-like state and is undocked and disordered in the ADP-bound structure (Fig. 2a). The same neck-linker positioning dependent on the switch II cluster was observed when all the plus-end-directed kinesin structures were superimposed (Fig. 2b). The switch-controlled movements of the neck linker complement data on nucleotide-dependent movements of the neck linker of conventional kinesin³⁵, and substantiate the proposal for how ATP-induced neck-linker docking can drive plus-end-directed kinesin motility^{34,35}.

In our study the conformational changes are seen for chimaeric protein, which has the neck of conventional kinesin attached to the catalytic core of KIF1A (Fig. 1a). Cooperative function between catalytic cores and exogenous neck domains has been reported previously for chimaeric kinesin proteins^{3,16,31–33}. The conserved kinesin catalytic core was suggested to function as an allosteric enzyme onto which class-specific neck domains serve as mechanical amplifiers and transmitters (transmissions and drive shafts) that confer unique motile properties to particular subfamilies of kinesin motors^{9,32,34–36}. Our study supports this hypothesis. We propose that the modular mechanism of nucleotide-driven conformational switching in the kinesin catalytic core can control actions of diverse kinesin mechanical elements and power motility of both plus- and minus-end-directed kinesin motors. The same conformational

change can be harnessed to produce forward or reverse movement. Because of the different location (N-terminal to the catalytic core) and different neck architecture of the minus-end-directed kinesin motors^{9,10} (Fig. 2c), the ADP-like position of the switch II cluster promotes docking of their neck to the core in the ADP state rather than undocking as in the forward-going kinesin. Moreover, the docked position of the neck in this case is stabilized by its specific interactions with the switch II cluster in the ADP state⁹ (Fig. 2c). With marked inverse symmetry, retraction of the switch II cluster in the ATP state of the minus-end-directed kinesins would disintegrate the neck/core interface and cause the neck to melt away from the core (a hypothetical transition of the neck of the minus-end-directed kinesin motor *ncd* is shown in Fig. 2c). This conformational transition would produce a force vector directed towards the microtubule minus-end, opposite to that of the conventional kinesin (Fig. 2a, c).

Switch-controlled affinity of the core for microtubules

The ultimate function of conformational switching in G proteins is to control their interactions with specific macromolecular targets—effectors^{17,37}. We propose that the conformational switching in kinesins evolved to control their interactions with microtubules. To compare interactions between the microtubule and the KIF1A catalytic core in ADP and ATP states, the structure of the KIF1A core bound to the microtubule in the presence of ADP was determined by cryo-electron microscopy (cryo-EM). The shape of the KIF1A–ADP core bound to the microtubule appears similar to that observed in the ATP-like state¹⁵ (Fig. 3a, c); however, the long axis of the core is rotated ~20 degrees clockwise in the ADP state (Fig. 3c) compared with the ATP-like state (Fig. 3a). Nucleotide-dependent conformational changes in the microtubule-bound kinesin heads observed in previous electron microscopy studies^{38–41} have been rather small. The large conformational change observed here might be explained by tighter binding of the KIF1A core to the microtubule in both nucleotide states (especially in the ADP state^{3,15,16}) and by enhanced resolution of the corresponding cryo-EM images.

To analyse the observed differences in positioning of the core on the microtubule, we docked the KIF1A and tubulin^{42,43} atomic models into the corresponding cryo-EM density maps using a real-space density-matching algorithm and manual fitting. Both

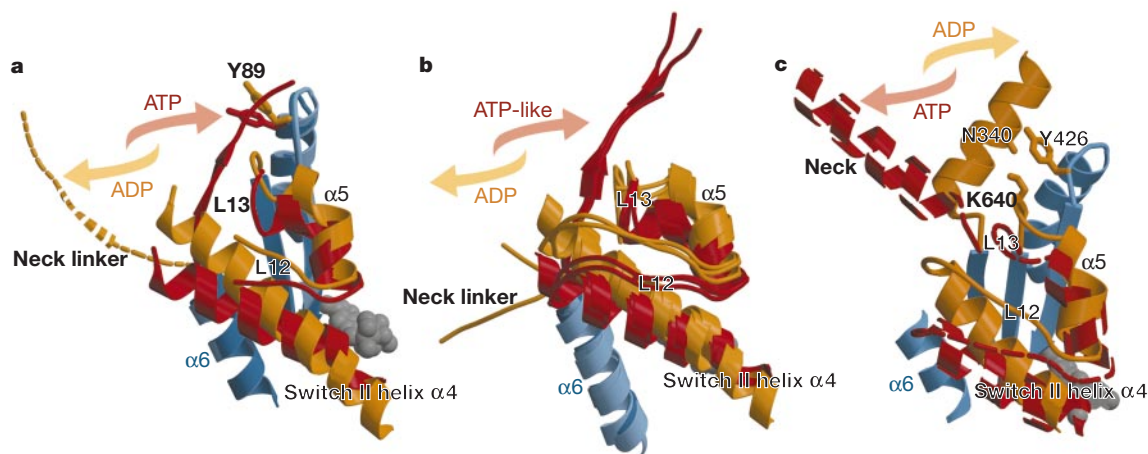


Figure 2 Nucleotide-dependent movements of the mechanical elements of kinesin motors. **a**, Reorientation of the neck linker between the ADP (yellow) and ATP-like (red) states of KIF1A. **b**, Superposition of the switch II clusters of the plus-end-directed kinesins. Conformation of the switch II cluster (ADP or ATP-like, yellow and red, respectively) controls position of the kinesin neck linker in all structures. Helix $\alpha 6$ is shown

in blue. **c**, Hypothetical model for the nucleotide-dependent dynamics at the core/neck interface of the minus-end-directed motor *ncd*⁹. The colours are the same as in **a** and **b**. The switch II cluster and the neck in ATP state are dashed. Conserved residues essential for stabilization of the neck/core interface in the ADP state of the *ncd*⁹ are labelled.

methods produced the same result: the long axis of the KIF1A atomic model in the ADP state was found rotated clockwise by ~ 20 degrees compared with its orientation in the ATP-like state (Fig. 3b, d).

Consistent with previous kinesin-docking experiments^{12,14,15}, the interactions between the KIF1A catalytic core and the microtubule are extensive ($\sim 900 \text{ \AA}^2$ in ADP and $\sim 1,200 \text{ \AA}^2$ in ATP-like state, without contribution of disordered loop L11 in either state) and involve three areas of the motor. The first at the plus-end side of the motor includes loop L8 (Fig. 4a, b, magenta). The second near the minus-end side of the motor includes helix $\alpha 6$ and possibly loop L2 (Fig. 4a, b, green). The switch II cluster in the centre of the core (Fig. 4a, yellow; Fig. 4b, red) forms the third contact area. Although most of the L11 and K-loop from the switch II cluster are disordered in the KIF1A crystal structures (Fig. 1b, c), these loops are ordered in the microtubule-bound motor (Fig. 4c). Notably, switch I loop L9 is also near the central microtubule-binding area (Fig. 4c).

The position of the switch II cluster is similar in ADP and ATP-like states of the microtubule-bound motor (Fig. 4d). This suggests that although in the unbound KIF1A structures the switch II cluster rotates by ~ 20 degrees relative to the core (Fig. 1b), when the motor

is bound to the microtubule the switch II cluster stays fixed to the microtubule, and the motor rotates by ~ 20 degrees in the opposite direction. This counter rotation of the motor could be explained by anchoring of the switch II cluster on the microtubule surface by numerous restraining binding contacts (Fig. 4a–c), which would prevent sliding of the cluster relative to the microtubule as ADP is replaced by ATP.

The nucleotide-dependent counter rotation causes a plus-end-directed displacement of the motor core on the microtubule by pointing the tip of the core (and the neck linker docked on the tip) towards the plus end of the microtubule in the ATP state (Fig. 4e). This plus-end-directed displacement might be further increased by sliding switch II helix $\alpha 4$ up the floor of its pocket on the microtubule surface (Fig. 4e). Because the floor of the microtubule pocket has a slope of ~ 45 degrees (Fig. 4a, b, e), sliding of $\alpha 4$ deeper inside the pocket in the ATP state results in a displacement of both the helix and the entire motor core towards the plus end of the microtubule. In the ADP state, $\alpha 6$ prevents this by steric interactions with the microtubule (Fig. 4a, e). By pointing the docked neck linker towards the plus end of the microtubule in the ATP state, the motor positions the C-terminal part of the linker in the vicinity of a distinct protrusion on the surface of the microtubule (E-hook¹⁶, Fig. 3b). According to the KIF1A–5-adenylylimidodiphosphate (AMPPNP)–microtubule model, this protrusion corresponds to flexible, negatively charged C-terminal regions of the upper tubulin subunit. Although absent in the atomic structure of the microtubule^{42,43}, this region seems to be partially structured in our EM-maps. Modelling experiments (data not shown) suggest that the region next to the KIF1A neck linker (the predicted helical part of the neck and/or the following fork-head-associated (FHA) domain, Fig. 1a) might interact with the C-terminal region of the upper tubulin subunit in the ATP state. Such interactions⁴⁴ would further increase the plus-end-directed bias of the KIF1A motor domain and could be important for the processivity of the motor (Fig. 5c). The plus-end-directed displacement of the microtubule-attached KIF1A core is analogous to that observed for other kinesins⁴¹. Therefore, the nucleotide-dependent reorientation of the core that we describe for KIF1A is probably the mechanism accounting for the plus-end-biased movement of kinesin cores^{16,36}.

The combined results of crystallographic and cryo-EM studies are sufficient for outlining a general switch-based mechanism for controlling affinity of the kinesin catalytic core for the microtubule. The conformational switching of the core between ADP and ATP states might control kinesin interactions with the microtubule in two ways. First, changes in molecular surface and electrostatic potential associated with the conformational transitions in the kinesin switch regions would influence affinity of the core for the microtubules in different nucleotide-bound states. The more convex shape of the switch II helix $\alpha 4$ in the ATP state (Figs 1b and 4e) could enable tighter contact between this helix and the complementary surface of the microtubule (Fig. 4e). This effect might be enhanced by both the extended binding contacts of the lengthened switch II loop L11 (Fig. 1c) and interactions of the switch I loop L9 (Fig. 4c) with the microtubule in the ATP state. The nucleotide-induced conformational changes in the switch regions affect the structure of yet another kinesin microtubule-binding site—loop L8, which is packed against and interacts directly with the switch I and switch II regions (Fig. 1b). Although the polypeptide backbone of L8 does not show large conformational changes between the ADP and ATP-like states of KIF1A, the rearrangement of the side chains in L8 is significant between the two states and probably effects binding of the catalytic core to the microtubules. The nucleotide-driven counter rotation of the core would provide the second, complementary way of controlling kinesin interactions with the microtubule by changing the accessibility of the polymer surface for the microtubule-binding sites of the motor. In particular, sliding of helix $\alpha 4$ deeper inside its pocket (Fig. 4e) would result in

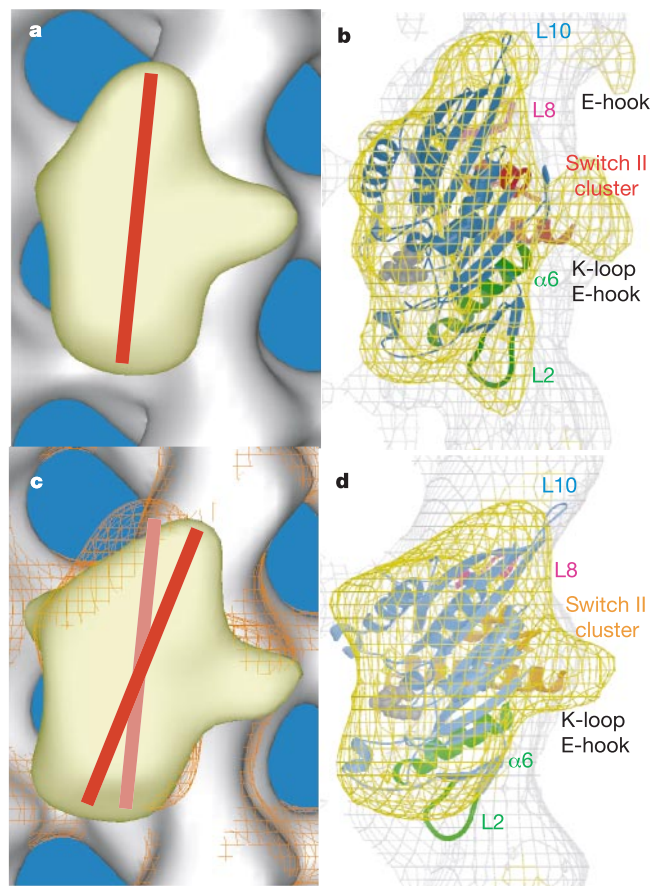


Figure 3 Cryo-EM maps of the microtubules decorated by the KIF1A. **a**, A 22 Å resolution map of the KIF1A–AMPPNP–microtubule complex¹⁵. The motor core (yellow) and its long axis (red line) are shown. The microtubules are shown with their plus end up in this and following figures. **b**, Docking of the ATP-like KIF1A crystal structure into 15 Å resolution cryo-EM map of the microtubule (grey) complexed with the KIF1A (yellow) in the presence of AMPPNP¹⁵. The C-terminal region of tubulin (E-hook) is shown in yellow. **c**, A 22 Å resolution map of the KIF1A–ADP–microtubule complex. The long axis of the motor is indicated by a red line. For comparison, its orientation in the ATP-like state is indicated by the orange grid and pink line. **d**, Docking of the ADP-bound KIF1A crystal structure into electron-microscopy-derived 22 Å resolution map. The colours are the same as in **b**.

higher affinity of the core for the microtubule in the ATP state.

As the nucleotide-dependent counter rotation of the core probably extends to all kinesins, it might constitute an important part of a general switch-based mechanism controlling actions of diverse kinesin motors (Fig. 5). In particular, it might be essential for the plus-end-biased brownian movement of the monomeric KIF1A motor^{3,16} (Fig. 5c). As shown in this (Figs 3 and 4) and previous studies^{3,15,16}, the positively charged K-loop of the KIF1A (Fig. 5c, green) can bind to the negatively charged flexible C-terminal regions of tubulin (E-hooks¹⁶, red in Fig. 5c). These electrostatic interactions might provide a flexible anchoring of the motor preventing its diffusion away from the microtubule while allowing diffusion along the microtubule in the weakly bound ADP state^{3,16}. In this state, the ridges that are formed between the tubulin subunits along the microtubule (Fig. 4a) might function as ratchets that capture the motor and minimize its diffusion towards the microtubule minus end. In the strongly bound ATP state, the nucleotide-induced counter rotation of the core would displace the motor towards the plus end of the microtubule, increasing the plus-end-directed bias of KIF1A movement. Furthermore, the counter rota-

tion of the core could induce transfer of the K-loop (arrow in Fig. 5c) from an E-hook at the minus-end side of the microtubule to the next one positioned towards the plus end of the microtubule. Docking the KIF1A neck linker in the ATP state might induce transfer of the FHA domain (arrow in Fig. 5c) from the minus-end side to the plus-end side of the motor, making the plus-end-biased processive movement of the KIF1A motor more efficient.

Regulation of conformational switching by microtubules

In all G proteins, the transitions between GDP and GTP states are controlled by specific regulators that accelerate nucleotide exchange (nucleotide exchange factors (NEFs)) or hydrolysis (GTPase-activating proteins (GAPs))^{17,37}. For kinesins, interactions with the microtubule accelerate both the nucleotide exchange and bond cleavage at different stages of the kinesin enzymatic cycle^{45,46}. Therefore, the microtubule might function not only as the analogue of the effector for G proteins, but may also combine roles of both the NEF and NTPase-activating protein¹⁸ for kinesins.

In G proteins complexed with NEFs, the tight interaction with GDP can be disrupted by displacing the switch I loop^{17,26}, by

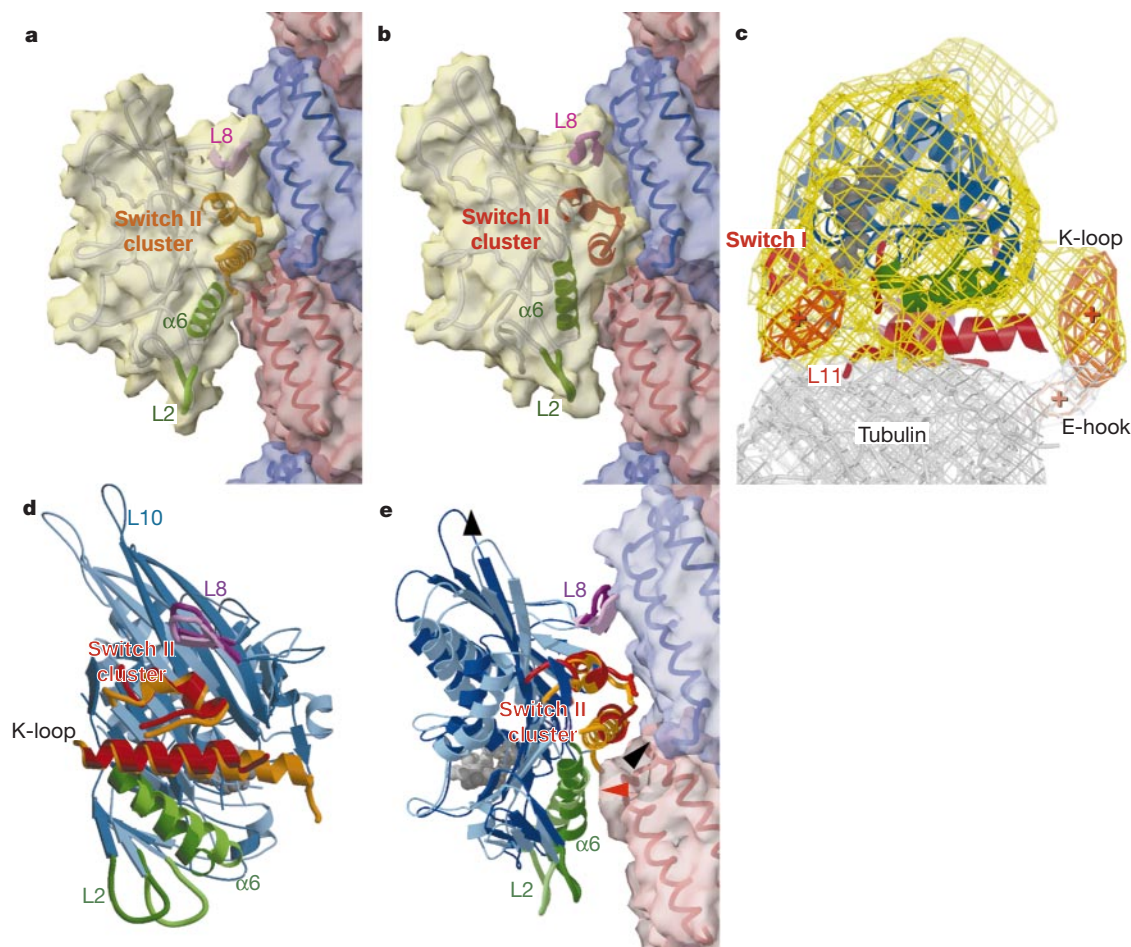


Figure 4 Interactions of KIF1A with the microtubule. **a, b**, Side views (~90 degrees relative to Fig. 3) of the docked atomic models of tubulin and KIF1A (**a**, ADP-bound; **b**, ATP-like). Molecular surfaces of the atomic models are coloured cream, blue and pink for the KIF1A, upper and lower tubulin subunits, respectively. The KIF1A microtubule-binding regions are highlighted and labelled. **c**, View (from the microtubule plus end) of the 15 Å resolution electron microscopy map (in yellow) of the KIF1A-AMPPNP-microtubule complex showing density (in red for the difference map between electron microscopy density and atomic model) corresponding to the structured loops L11 and K-loop. The

switch I loop L9 is in the vicinity of the microtubule. **d**, Tubulin-navigated superposition of the ADP and ATP-like KIF1A crystal structures from the KIF1A-ADP-microtubule and KIF1A-AMPPNP-microtubule complexes (view from the microtubule). The colours are the same as in Fig. 2. In addition, the regions involved in binding with the microtubule are indicated. **e**, Tubulin-navigated superposition of the ADP and ATP-like KIF1A structures docked on the microtubule. The plus-end-directed displacement of the motor in the ATP-like state and steric contact of $\alpha 6$ with the microtubule in the ADP state are indicated by black and red arrowheads, respectively.

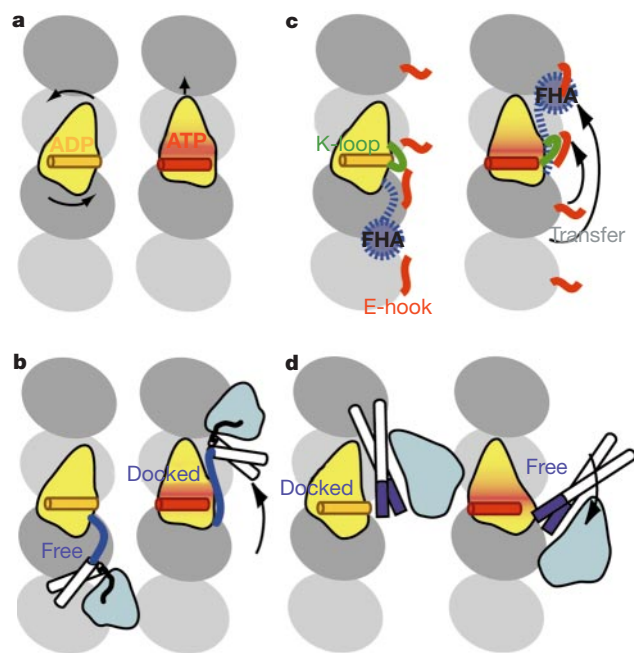


Figure 5 Models for motility by various kinesin motors. **a**, The nucleotide-induced counter rotation of the kinesin core (yellow) provides a unifying mechanism for different kinesins. It results in tighter binding (pink zone) and the plus-end-directed displacement of the core (shown by arrowhead) in the ATP state. The switch II cluster is shown as yellow and red for the ADP and ATP states, respectively. **b**, In the conventional dimeric kinesin, the plus-end-directed movement of the core is amplified by the nucleotide-controlled actions of the neck linker (purple) and the second head (blue) of the motor²⁵. **c**, In the monomeric kinesin motor KIF1A, actions of the class-specific K-loop (green), the neck region and FHA domain (both purple) might complement movements of the core and allow the plus-end-biased movement of the motor. The C-terminal regions of tubulin (E-hooks) are in red. **d**, The reversed mechanics linking the neck and the switch II cluster of the minus-end-directed kinesins (see Fig. 3c) might explain the opposite polarity of movement of these motors.

distorting the switch II region^{17,26} or by perturbing the binding site for the nucleotide base^{17,47}. Any of these mechanisms could be used for the release of tightly bound ADP from the kinesins. Displacement of the switch I region is probable considering the flexible character of the switch I loop and its proximity to the microtubule surface in the KIF1A–ADP–microtubule complex. Furthermore, as the switch II region constitutes a part of the binding interface in the ADP state of the microtubule-bound motor (Fig. 4a), its structure and, therefore, the binding site for the phosphate groups and magnesium ion could be affected by the motor's interaction with the microtubule. Finally, binding contacts with the microtubule at the minus-end side of the motor ($\alpha 6$ and possibly L2, Fig. 4a) could displace residues that bind the nucleotide base. The latter mechanism might relate the presence of the class-specific insertion in loop L3 of the monomeric kinesins with the higher nucleotide exchange rate of KIF1A (ref. 3).

In all G protein–GAP complexes, GAPs interact with both of the switch regions, assuring the precise positioning of residues involved in GTP hydrolysis^{17,25,37}. Analogous to G proteins, in the ATP state of kinesins, structuring of both of the switch regions by contacts with the microtubule might be crucial for proper positioning of residues involved in ATP hydrolysis. In kinesins, switch I serine and switch II glycine (S215 and G251 in KIF1A, Fig. 1c) coordinate the γ -phosphate of the nucleotide. The current model of the ATP-like KIF1A bound to the microtubule (Fig. 4a, c) shows the neighbouring residues, conserved R216 of switch I and E253 and N272 of

switch II, in the vicinity of the microtubule. These three side chains might mediate microtubule-assisted structuring of the kinesin catalytic site, consistent with results of mutational studies^{35,48}. Previous structural analysis of both kinesin and myosin motors showed that the switch I arginine and switch II glutamate (R216 and E253 in KIF1A) could tether by means of a salt bridge in the ATP state^{18,22,24}. The ATP-like structure of KIF1A shows these side chains untethered. Further structural studies at atomic resolution are needed to relate the kinesin active site and the molecular mechanism of the microtubule-assisted ATP hydrolysis. □

Methods

Expression, purification and crystallization of KIF1A

Because of difficulties in expression of the KIF1A motor domain, we instead chose the chimeric KIF1A construct^{3,15}. Crystallization was further aided by deleting 11 C-terminal residues from the construct C351 (refs. 3, 15, 16). To produce the construct C340, a DNA fragment encoding residues 1–355 of the KIF1A followed by residues 335–340 of a murine conventional kinesin KIF5C and (His)₆ tag was synthesized by PCR, cloned into the vector pET21b (Novagen) and transformed into BL21 (DE3) cells (Novagen) for expression. Cells were grown for 10 h at 24 °C in $\times 3$ YT media with ampicillin (0.1 mg ml^{−1}), packed, resuspended in buffer A (10 mM Tris–HCl pH 8.5, 500 mM NaCl, 60 mM imidazole and protease inhibitors) and lysed using a French pressure cell press (American Instrument). Soluble protein fraction was loaded onto a charged chelating–Sephacose Fast Flow (Pharmacia) column equilibrated with buffer A, and 500 mM of imidazole was used to elute the protein. The pooled fractions were dialysed against buffer B (10 mM MOPS pH 7.0, 100 mM NaCl, 1 mM EGTA, 0.1 mM ADP, 1 mM MgCl₂ and 20% w/v sucrose), loaded onto a RESOURCE S (Akta) column equilibrated with buffer C (100 mM MES–NaOH pH 6.0, 15% w/v sucrose, 1 mM EGTA, 1 mM MgCl₂), and a linear gradient of NaCl (0.1–0.5 M) was applied to elute the protein. The pooled protein fractions were dialysed against buffer B, concentrated to ~ 15 mg ml^{−1} and used for crystallization by vapour diffusion method. For KIF1A–ADP crystals, 2 μ l of protein solution were mixed with 2 μ l of reservoir buffer (RB1) (30% w/v PEG4000, 100 mM Tris–HCl pH 8.5, 200 mM sodium acetate, 8% w/v sucrose, 4 mM ADP, 10 mM MgCl₂) and equilibrated against RB1 for 5 d.

Crystals of the ATP-like KIF1A were obtained using RB2 (27% w/v PEG4000, 100 mM MES–NaOH pH 6.5, 200 mM sodium acetate) and then soaked in RB2 with 20 mM AMPPCP and 40 mM MgCl₂ for 24 h before data collection. Similar crystals were obtained by co-crystallizing the protein sample with 5 mM AMPPCP and 20 mM MgCl₂ in buffer RB3 (30% w/v PEG4000, 100 mM Tris–HCl pH 8.5, 200 mM sodium acetate). The ATP-like conformation was initially observed for the KIF1A crystallized in RB2 in the absence of AMPPCP. This result prompted studies on KIF1A in the presence of different nucleotide analogues.

Data collection, model building and refinement

We measured X-ray diffraction data at -180 °C using mother liquor as cryo-solvent. For the ADP-bound crystal, the data were collected to 2.2 Å at Stanford Synchrotron Radiation Laboratory beamline 9-1 (wavelength $\lambda = 0.98$ Å). For the AMPPCP-bound crystals, the data were collected to 2.0 and to 1.9 Å (with soaked-in and co-crystallized AMPPCP, respectively) at Advanced Light Source (Lawrence Berkeley National Laboratory) beamline 5.0.2 ($\lambda = 1.1$ Å). All data sets were integrated using DENZO and scaled with SCALEPACK⁴⁹. All crystal forms were of the orthorhombic space group $P2_12_12_1$ with one KIF1A–nucleotide complex in the asymmetric unit and cell dimensions of $a = 41.67$ Å, $b = 51.93$ Å, $c = 157.06$ Å; $a = 41.98$ Å, $b = 56.40$ Å, $c = 156.12$ Å; and $a = 42.42$ Å, $b = 55.43$ Å, $c = 157.27$ Å for ADP, AMPPCP-soaked and AMPPCP co-crystallized forms, respectively.

The AMPPCP-bound KIF1A structures were determined by molecular replacement method (CNS⁵⁰) using atomic coordinates for the conventional rat kinesin⁷. Electron-density maps based on coefficients $2F_o - F_c$ were calculated from the phases of the initial model. Subsequent rounds of model building and refinement were performed using programs QUANTA (Molecular Simulations) and CNS⁵⁰, respectively. Electron density corresponding to the bound AMPPCP was well defined (see Supplementary Information) in maps calculated for the KIF1A structures with both soaked-in and co-crystallized AMPPCP. Both structures were virtually identical; however, L9 was better ordered in the KIF1A structure with AMPPCP soaked in. The latter structure is refined to R and R_{free} values of 21.9 and 23.0, respectively (20.0–2.0 Å) and includes 328 KIF1A amino-acid residues together with Mg–AMPPCP complex, 230 water molecules and 5 residues (362–366) from the (His)₆ tag. Residues 1–2, 254–268, 290–302, 359–361 and 367 are disordered.

The ADP-bound structure of KIF1A was determined by molecular replacement method (CNS⁵⁰) using atomic coordinates for the AMPPCP-bound KIF1A. We performed model building and refinement as described above. The model is refined to R and R_{free} values of 21.0 and 22.5, respectively (20.0–2.2 Å) and includes 330 amino-acid residues together with Mg–ADP complex and 235 water molecules. Residues 1–3, 256–260, 290–303, and 353–367 are disordered.

Electron microscopy and docking of atomic models

A sample of microtubules decorated by C351 in ADP state was prepared as described¹⁵ using 50 mM imidazole, 5 mM Mg–acetate, 1 mM EGTA, 50 mM potassium acetate, 10 mM dithiothreitol, 0.1% Triton X-100, 2 mM ADP and 10 μ M paclitaxel. Cryo-EM

images were taken with the defocus range of 17,000–33,000 Å and processed as described¹⁵. The final map was calculated by combining the best 90 data sets that represent ~13,000 asymmetric units. The effective resolution (22 Å, see Supplementary Information) was determined by Fourier shell correlation between two independently averaged data sets. Docking of the atomic models was performed using a real-space density matching algorithm used in EOS system (Yasunaga, T. and Wakabayashi, T. (Univ. Tokyo) with a difference that the Fast Fourier Transform method was used for translational search.

Received 29 December 2000; accepted 19 April 2001.

1. Okada, Y., Yamazaki, H., Sekine-Aizawa, Y. & Hirokawa, N. The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell* **81**, 769–780 (1995).
2. Hirokawa, N. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* **279**, 519–526 (1998).
3. Okada, Y. & Hirokawa, N. A processive single-headed motor: kinesin superfamily protein KIF1A. *Science* **283**, 152–157 (1999).
4. Kull, F. J., Sablin, E. P., Lau, R., Fletterick, R. J. & Vale, R. D. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature* **380**, 550–555 (1996).
5. Sablin, E. P., Kull, F. J., Cooke, R., Vale, R. D. & Fletterick, R. J. Crystal structure of the motor domain of the kinesin-related motor ncd. *Nature* **380**, 555–559 (1996).
6. Sack, S. *et al.* X-ray structure of motor and neck domains from rat brain kinesin. *Biochemistry* **36**, 16155–16165 (1997).
7. Kozielski, F. *et al.* The crystal structure of dimeric kinesin and implications for microtubule-dependent motility. *Cell* **91**, 985–994 (1997).
8. Gulick, A., Song, H., Endow, S. & Rayment, I. X-ray crystal structure of the yeast KAR3 motor domain complexed with MgADP to 2.3 Å resolution. *Biochemistry* **37**, 1769–1776 (1998).
9. Sablin, E. P. *et al.* Direction determination in the minus-end-directed kinesin motor ncd. *Nature* **395**, 813–816 (1998).
10. Kozielski, F., De Bonis, S., Burmeister, W. P., Cohen-Addad, C. & Wade, R. H. The crystal structure of the minus-end-directed microtubule motor protein ncd reveals variable dimer conformations. *Struc. Fold Des.* **7**, 1407–1416 (1999).
11. Woehlke, G. *et al.* Microtubule interaction site of the kinesin motor. *Cell* **90**, 207–216 (1997).
12. Sosa, H. *et al.* A model for the microtubule-ncd motor protein complex obtained by cryo-electron microscopy and image analysis. *Cell* **90**, 217–224 (1997).
13. Alonso, M., Damme, J., Vandekerckhove, J. & Cross, R. A. Proteolytic mapping of kinesin/ncd-microtubule interface: nucleotide-dependent conformational changes in the loops L8 and L12. *EMBO J.* **17**, 945–951 (1998).
14. Hoenger, A. *et al.* Image reconstructions of microtubules decorated with monomeric and dimeric kinesins: comparison with X-ray structure and implications for motility. *J. Cell Biol.* **141**, 419–430 (1998).
15. Kikkawa, M., Okada, Y. & Hirokawa, N. 15 Å resolution model of the monomeric kinesin motor, KIF1A. *Cell* **100**, 241–252 (2000).
16. Okada, Y. & Hirokawa, N. Mechanism of the single-headed processivity: diffusional anchoring between the K-loop of kinesin and the C terminus of tubulin. *Proc. Natl Acad. Sci. USA* **97**, 640–645 (2000).
17. Sprang, S. R. G protein mechanism: insights from structural analysis. *Annu. Rev. Biochem.* **66**, 639–678 (1997).
18. Vale, R. D. Switches, latches, and amplifiers: common themes of molecular motors and G proteins. *J. Cell Biol.* **135**, 291–302 (1996).
19. Rayment, I. *et al.* Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–58 (1993).
20. Dominguez, R., Frey, Y., Trybus, K. M. & Cohen, C. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. *Cell* **94**, 559–571 (1998).
21. Houdusse, A., Kalabokis, V. N., Himmel, D., Szent-Gyorgyi, A. G. & Cohen, C. Atomic structure of scallop myosin subfragment S1 complexed with MgADP: a novel conformation of the myosin head. *Cell* **97**, 459–470 (1999).
22. Geeves, M. A. & Holmes, K. C. Structural mechanism of muscle contraction. *Ann. Rev. Biochem.* **68**, 687–728 (1999).
23. Milburn, M. V. *et al.* Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247**, 939–945 (1990).
24. Sack, S., Kull, F. J. & Mandelkow, E. Motor proteins of the kinesin family: structures, variations, and the nucleotide binding sites. *Eur. J. Biochem.* **262**, 1–11 (1999).
25. Scheffzek, K. *et al.* The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338 (1997).
26. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. The structural basis of the activation of Ras by Sos. *Nature* **394**, 337–343 (1998).
27. Kjeldgaard, M., Nyborg, J. & Clark, B. F. The GTP binding motif: variations on a theme. *FASEB J.* **10**, 1347–1368 (1996).
28. Vale, R. D. & Milligan, R. A. The way things move: looking under the hood of molecular motor proteins. *Science* **288**, 88–95 (2000).
29. Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. The crystal structure of elongation factor Ef-Tu from *Thermus aquaticus* in the GTP conformation. *Structure* **1**, 35–50 (1993).
30. Polekhina, G., Thirup, S., Kjeldgaard, M., Nissen, P. & Lippmann, C. Helix unwinding in the effector region of elongation factor Ef-Tu-GDP. *Structure* **4**, 1141–1151 (1996).
31. Henningsen, U. & Schliwa, M. Reversal in the direction of movement of a molecular motor. *Nature* **389**, 93–95 (1997).
32. Case, R. B., Pierce, D. W., Hom-Booher, N., Hart, C. & Vale, R. D. The directional preference of kinesin motors is specified by an element outside of the motor catalytic domain. *Cell* **90**, 959–966 (1997).
33. Endow, S. A. & Waligora, K. W. Determinants of kinesin motor polarity. *Science* **281**, 1200–1202 (1998).
34. Vale, R. D., Case, R., Sablin, E., Hart, C. & Fletterick, R. Searching for kinesin's mechanical amplifier. *Phil. Trans. R. Soc. Lond. B* **355**, 449–457 (2000).
35. Rice, S. *et al.* A structural change in the kinesin motor protein that drives motility. *Nature* **402**, 778–784 (1999).
36. Case, R. B., Rice, S., Hart, C. L., Ly, B. & Vale, R. D. Role of the kinesin neck linker and catalytic core in microtubule-based motility. *Curr. Biol.* **10**, 157–160 (2000).
37. Bohm, A., Gaudet, R. & Sigler, P. Structural aspects of heterotrimeric G-protein signalling. *Curr. Opin. Biotech.* **8**, 480–487 (1997).
38. Hirose, K., Lockhart, A., Cross, R. A. & Amos, L. A. Nucleotide-dependent angular change in kinesin motor domain bound to tubulin. *Nature* **376**, 277–279 (1995).
39. Hirose, K., Cross, R. A. & Amos, L. A. Nucleotide-dependent structural changes in dimeric ncd molecules complexed to microtubules. *J. Mol. Biol.* **278**, 389–400 (1998).
40. Arnal, I. & Wade, R. H. Nucleotide-dependent conformations of the kinesin dimer interacting with microtubules. *Structure* **6**, 33–38 (1998).
41. Hirose, K. *et al.* Structural comparison of dimeric Eg5, *Neurospora* kinesin (Nkin) and NCD head-Nkin neck chimera with conventional kinesin. *EMBO J.* **19**, 5308–5314 (2000).
42. Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature* **391**, 199–203 (1998).
43. Nogales, E., Whittaker, M., Milligan, R. A. & Downing, K. H. High-resolution model of the microtubule. *Cell* **96**, 79–88 (1999).
44. Thorn, K. S., Ubersax, J. A. & Vale, R. D. Engineering the processive run length of the kinesin motor. *J. Cell Biol.* **151**, 1093–1100 (2000).
45. Ma, Y.-Z. & Taylor, E. W. Interacting head mechanism of microtubule-kinesin ATPase. *J. Biol. Chem.* **272**, 724–730 (1997).
46. Gilbert, S. P., Moyer, M. L. & Johnson, K. A. Alternating site mechanism of the kinesin ATPase. *Biochemistry* **37**, 792–799 (1998).
47. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusak, S. & Leberman, R. The structure of the *Escherichia coli* EF-Tu-EF-Ts complex at 2.5 Å resolution. *Nature* **379**, 511–518 (1996).
48. Song, H. & Endow, S. A. Decoupling of nucleotide- and microtubule-binding sites in a kinesin mutant. *Nature* **396**, 587–590 (1998).
49. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
50. Brunger, A. T. *et al.* Crystallography and NMR system: a new software suit for macromolecular structure determination. *Acta Cryst. D* **54**, 905–921 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank C. Sindelar and J. Turner for providing comments on this manuscript and sharing their results before publication. We also thank M. Sugaya for technical assistance. This work was supported by a Center of Excellence Grant-in-aid from the Ministry of Education, Science, Sports, and Culture of Japan (N.H.), and by funding from an NIH program project grant.

Correspondence and requests for materials should be addressed to N.H. (e-mail: hirokawa@m.u-tokyo.ac.jp). Coordinates have been deposited in the Protein Data Bank under numbers 1I61, 1I55 and 1IA0.

The size and albedo of the Kuiper-belt object (20000) Varuna

David Jewitt*, Herve Aussel* & Aaron Evans†

* Institute for Astronomy, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Department of Physics and Astronomy, SUNY at Stony Brook, Stony Brook, New York 11794-3800, USA

Observations over the last decade have revealed the existence of a large number of bodies orbiting the Sun beyond Neptune¹. Known as the Kuiper-belt objects (KBOs), they are believed to be formed in the outer reaches of the protoplanetary disk around the young Sun, and have been little altered since then. They are probably the source of short-period comets². The KBOs are, however, difficult objects to study because of their distance from earth, so even basic physical properties such as their sizes and albedos remain unknown. Previous size estimates came from assuming an albedo with the canonical value being 0.04. Here we report simultaneous measurements of the thermal emission and reflected optical light of the bright KBO (20000) Varuna, which allow us to determine independently both the size and the albedo. Varuna has an equivalent circular diameter of $D = 900^{+129}_{-145}$ km and a red geometric albedo of $p_R = 0.070^{+0.030}_{-0.017}$. Its surface is darker than Pluto's, suggesting that it is largely devoid of fresh ice, but brighter than previously assumed for KBOs.

The most fundamental physical properties of KBOs are their size and albedo. All published estimates of the sizes of KBOs are on the basis of an assumed albedo of 0.04—the albedo of the cometary nuclei^{3–6}—and may be considerably in error¹. Optical observations provide only a measure of the product of the geometric albedo with the physical cross-section. The albedo and cross-section can be determined separately by combining optical data with a simultaneous measurement of the thermally emitted flux⁷. The thermal emission results from surface heating by sunlight, and varies in proportion to $(1 - \text{albedo})$ times the cross-section. The two constraints supplied by the optical and thermal wavelength observations permit a solution for the two principal unknowns, the albedo and the size. This technique has been applied widely to main-belt asteroids, for which the Planck maximum falls in the observationally accessible 10–20 μm wavelength range. At heliocentric distance $R = 40$ AU, however, the KBOs have equilibrium blackbody temperatures near 45 K and the Planck maximum at 70 μm corresponds to a wavelength at which the atmosphere of the Earth is opaque. Therefore, measurements of thermal emission from KBOs must be taken from space, or through submillimetre windows in the atmospheric transmission. No ground-based thermal observations have been obtained before now owing to the faintness of KBO targets in the atmospheric windows.

The optically bright KBO (20000) Varuna (formerly 2000 WR106) was discovered on 28 November 2000 by the Spacewatch telescope⁸. Observations before discovery dating back to 1954 allowed a quick classification as a classical KBO (ref. 1) with semi-major axis 43.27 AU⁹. Submillimetre observations of (20000) Varuna were obtained using the Submillimeter Common User Bolometer Array (SCUBA) at the James Clerk Maxwell Telescope (JCMT) on Mauna Kea, Hawaii, on UT 30–31 December 2000 with simultaneous optical photometry from the University of Hawaii 2.2-m telescope. At the time of observation, the heliocentric and geocentric distances of (20000) Varuna were $R = 43.05$ AU and $\Delta = 42.06$ AU, respectively. The phase angle was 0.03°. See Table 1 for the photometric measurements.

The density of faint background galaxy sources brighter than 3 mJy at 850 μm is about 10^3 deg^{-2} (ref. 10). The probability that such a source might fall in the SCUBA beam is about 0.01. As a precaution

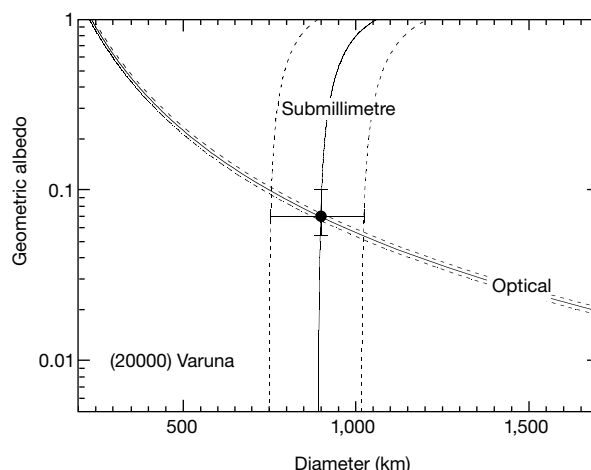


Figure 1 Albedo and diameter constraints from the optical and submillimetre thermal observations of (20000) Varuna.

against this (small) possibility, we took data on two nights so that the motion of (20000) Varuna would prevent any single background object from contaminating the photometry. The probability that two different background objects might contaminate the SCUBA beam on two nights is of the order of 10^{-4} and can therefore be neglected. In addition, we find no galactic source in the Infrared Astronomical Satellite 60- μm survey within 5 arcmin of (20000) Varuna.

The weak temperature dependence of the Planck function in the Rayleigh–Jeans limit conveys a significant advantage to the interpretation of submillimetre thermal emission data¹¹. Specifically, uncertainties in the spatial distribution of the surface temperature¹² have a minimal effect on the Rayleigh–Jeans emission, whereas their impact on the flux emitted near the Planck maximum can be very strong^{13,14}. In the Rayleigh–Jeans limit, the diameter is given by¹¹:

$$D = \lambda \Delta \left(\frac{2S_v}{\pi k \epsilon_{\text{sm}}} \right)^{1/2} \left(\frac{\sigma R^2}{F_{\text{Sun}}} \right)^{1/8} \left(\frac{\epsilon_{\text{ir}} \chi}{1 - A} \right)^{1/8} \quad (1)$$

where $\lambda = 850 \mu\text{m}$ is the wavelength of observation, Δ (m) is the geocentric distance, S_v ($\text{W m}^{-2} \text{ Hz}^{-1}$) is the measured flux density, $k = 1.38 \times 10^{-23} \text{ J K}^{-1}$ is Boltzmann's constant, $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ is the Stefan–Boltzmann constant, R (AU) is heliocentric distance, and $F_{\text{Sun}} = 1,360 \text{ W m}^{-2}$ is the solar constant. The emissivity of the surface at submillimetre and thermal infrared (Planck maximum) wavelengths is given by ϵ_{sm} and ϵ_{ir} respectively. We assume $\epsilon_{\text{sm}} = \epsilon_{\text{ir}} = 0.9$. The factor χ parametrizes the distribution of temperature over the surface where $\chi = 2$ corresponds to the 'standard thermal model' approximation¹⁵ in which the Sun is on the rotation axis and solar heating is confined to a single hemisphere, whereas $\chi = 4$ corresponds to an isothermal body. The intermediate case, a body illuminated equatorially and having a short rotation period compared with the cooling time (the so-called 'isothermal latitude model') has $\chi = \pi$. We assume $\chi = 2$. Last in equation (1), A is the (dimensionless) bond albedo, related to the geometric albedo by:

$$A = p_R q \quad (2)$$

where q is the 'phase integral' and p_R is the R-band geometric albedo. Measurements of airless planetary bodies suggest $q = 0.75$, with an uncertainty of about 30%¹⁶. In equation (1) even factor-of-two uncertainties in χ , ϵ_{ir} and $(1 - A)$ correspond to only 10% uncertainties in the derived diameter, owing to the one-eighth power dependence. Equation (1) is more strongly dependent on ϵ_{sm} . Measurements of Ganymede, Callisto and Pluto all suggest $\epsilon_{\text{sm}} \approx 0.9$, as used here¹¹. A reduced submillimetre emissivity by a

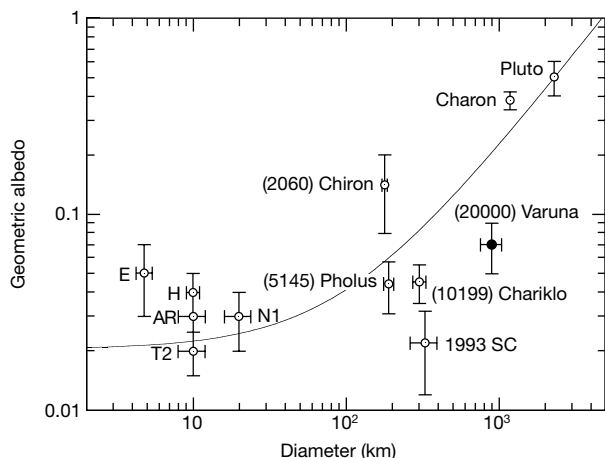


Figure 2 Well determined albedos and diameters of outer Solar System objects. The nuclei of short-period comets are: E, P/Encke⁶; AR, P/Arend-Rigaux⁵; T2, P/Tempel 2 (ref. 4); N1, P/Neujmin 1 (ref. 3); and H, P/Halley²². Other objects include centaurs (2060) Chiron²³, (5145) Pholus²⁴ and (10199) Chariklo¹³, and Pluto and Charon²⁵. The albedo of Charon should be considered an upper limit, owing to near-nucleus dust coma contamination. A 2.7σ confidence limit to the 90- μ m flux from the Infrared Space Observatory on 1993 SC was used to obtain estimates for the diameter (328^{+58}_{-66} km) and albedo ($0.022^{+0.013}_{-0.006}$), while object 1996 TL66 was not detected at its expected position¹⁴. Thermal observations of KBOs will be an important objective of NASA's upcoming SIRTf satellite. The solid line shows the linear correlation between albedo and diameter. The linear correlation coefficient for the objects is 0.92, and the probability that this or a larger correlation might be obtained by chance from uncorrelated data is $P(0.92, 12) < 10^{-4}$. Removal of Pluto from the plot reduces the correlation coefficient to 0.79, for which $P(0.79, 11) = 0.004$, which is statistically insignificant (3σ corresponds to $P = 0.003$). Plotted error bars are all 1σ . $P(x, y)$ notation indicates the correlation coefficient, x , and the number of data points, y .

factor of two would increase the derived diameter by $2^{1/2}$ and decrease the albedo by a factor of 2; however, we consider it unlikely that ϵ_{sm} is different from 0.9 by a factor of 2.

The apparent red magnitude, m_R , of a body viewed in scattered sunlight is related to its physical properties by¹⁷:

$$\phi(\alpha)p_R D^2 = 8.96 \times 10^{22} R^2 \Delta^{10} 10^{0.4(m_{\text{Sun}} - m_R)} \quad (3)$$

where $\phi(\alpha)$ is the phase function at phase angle α (degrees), D (m) is the diameter, R (AU) and Δ (AU) the heliocentric and geocentric distances, respectively, and $m_{\text{Sun}} = -27.1$ is the red magnitude of the Sun. The phase angle is small so that we may set $\phi(0.03^\circ) = 1$. The optical photometry provides, through equation (3), a constraint on the product $p_R D^2$.

The trajectories defined by equations (1) and (2) are plotted in Fig. 1. The trajectories intersect at $D = 900^{+125}_{-145}$ km, $p_R = 0.070^{+0.030}_{-0.017}$, which we take as the best estimate of the diameter and albedo of (20000) Varuna. The quoted uncertainties reflect only statistical errors. The optical brightness of (20000) Varuna varies by 0.5 mag (50%) with a period near 3 h (ref. 18), and the phase of the optical light curve at the time of our observations is unknown. This introduces a systematic uncertainty into the derived albedo of up to 50%. A further uncertainty of the order of 10% is introduced by the small phase angle at the epoch of observation (Table 1) and by the unknown magnitude of the opposition surge¹⁹.

Figure 2 compares the albedos and diameters of a number of well studied bodies in the outer Solar System. The observations show a trend towards higher albedos at larger sizes. The small cometary nuclei are the darkest of the observed bodies, whereas Pluto and its satellite Charon are the most reflective. This trend is highly statistically significant but disappears when Pluto is removed from the sample (Fig. 2, legend). The globally high albedo of Pluto is the result of surface frosts freshly deposited from a tenuous

Table 1 Photometry of (20000) Varuna

Date (UT)	Telescope	Wavelength	Flux density/magnitude
2000 December 30.4764–30.5625	JCMT	850 μ m	3.14 ± 1.10 mJy
2000 December 31.4896–31.5764	JCMT	850 μ m	2.37 ± 1.29 mJy
Combined		850 μ m	2.81 ± 0.85 mJy
2000 December 30.4757	UH 2.2 m	0.65 μ m	19.7 ± 0.1 mag

The JCMT provides diffraction-limited (14 arcsec full width at half maximum (FWHM)) resolution images at $\lambda = 850 \mu$ m. We used the SCUBA camera²⁰ in photometry mode with hourly pointing and photometric calibration obtained from nearby fixed objects CRL618, 0.745+241 and OJ287. The pointing was accurate to ± 2 arcsec, while the ephemeris accuracy was better than 1 arcsec. The telescope was tracked on the motion of (20000) Varuna, at about 3.1 arcsec h⁻¹ westward, using software to interpolate the motion. Atmospheric opacity was monitored using skydips made with the JCMT and with a radiometer at the nearby Caltech Submillimetre Observatory. The 230 GHz optical depth of the atmosphere varied in the range 0.065–0.080, corresponding to a water column abundance of about 1.5 mm. Sky cancellation was achieved by chopping to a position 90 arcsec from the source in elevation at a frequency of 7.8 Hz. Accumulated observations of about 5 h were secured on (20000) Varuna. Optical observations were taken at the UH 2.2-m telescope, in the period UT 30.4715–30.5368 December 2000. The 'Orbit' 2048 $\times$ 2048 pixel charge-coupled device was used, with image scale 0.28 arcsec per binned pixel. The flat-fielded red (0.65 μ m) filter images were photometrically calibrated using observations of stars from ref. 21. Integration time was 180 s.

atmosphere. The smaller (20000) Varuna, with a radius 40% that of Pluto and perhaps only 6% of Pluto's mass, has insufficient gravity to retain an atmosphere, and consequently lacks a global surface frost. Patchy frost distributions are not excluded, however, and would be compatible with the recently measured large, rotational lightcurve¹⁸. □

Received 18 January; accepted 27 March 2001.

- Jewitt, D. & Luu, J. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. & Russell, S.) 1201–1229 (Univ. Arizona Press, Tucson, 2000).
- Duncan, M., Quinn, T. & Tremaine, S. The origin of short-period comets. *Astrophys. J.* **328**, L69–L73 (1988).
- Campins, H., A'Hearn, M. & McFadden, L. The bare nucleus of comet Neujmin 1. *Astrophys. J.* **316**, 847–857 (1987).
- A'Hearn, M., Campins, H., Schleicher, D. & Millis, R. The nucleus of Comet P/Tempel 2. *Astrophys. J.* **347**, 1155–1166 (1989).
- Millis, R., A'Hearn, M. & Campins, H. An investigation of the nucleus and coma of Comet P/Arend-Rigaux. *Astrophys. J.* **324**, 1194–1209 (1988).
- Fernandez, Y. et al. Physical properties of the nucleus of comet 2P/Encke. *Icarus* **147**, 145–160 (2000).
- Allen, D. A. Infrared diameter of Vesta. *Nature* **227**, 158–159 (1970).
- McMillan, R. & Larsen, J. 2000 WR106. *Minor Planet Electronic Circular* 2000-X02 (cfa-www.harvard.edu/iau/mpec/K00/K00X02.html) (2000).
- Knofel, A. & Stoss, R. 2000 WR106. *Minor Planet Electronic Circular* 2000-Y45 (cfa-www.harvard.edu/iau/mpec/K00/K00Y45.html) (2001).
- Barger, A., Cowie, L. & Sanders, D. Resolving the submillimeter background: the 850 micron Galaxy counts. *Astrophys. J.* **518**, L5–L8 (1999).
- Jewitt, D. C. & Luu, J. X. Submillimeter continuum observations of 2060 Chiron. *Astron. J.* **104**, 398–404 (1992).
- Spencer, J. R., Lebofsky, L. A. & Sykes, M. V. Systematic biases in radiometric diameter determinations. *Icarus* **78**, 337–354 (1989).
- Jewitt, D. C. & Kalas, P. Thermal observations of Centaur 1997 CU26. *Astrophys. J.* **499**, L103–L106 (1998).
- Thomas, N. et al. Observations of the trans-Neptunian objects 1993 SC and 1996 TL66 with the infrared space observatory. *Astrophys. J.* **534**, 446–455 (2000).
- Lebofsky, L. et al. A refined 'standard' thermal model for asteroids based on observations of 1 Ceres and 2 Pallas. *Icarus* **68**, 239–251 (1986).
- Harris, D. L. in *Planets and Satellites* (eds Kuiper, G. & Middlehurst, B.) 272–342 (Univ. Chicago Press, Chicago, 1961).
- Russell, H. N. On the albedo of the planets and their satellites. *Astrophys. J.* **43**, 173–195 (1916).
- Farnham, T. (20000) 2000 WR 106. *IAU Circular* 7583 (2001).
- Bowell, E. et al. in *Asteroids II* (eds Binzel, R., Gehrels, T. & Matthews, M.) 524–556 (Univ. Arizona Press, Tucson, 1989).
- Holland, W. S. et al. SCUBA: a common-user submillimetre camera operating on the James Clerk Maxwell Telescope. *Mon. Not. R. Astron. Soc.* **303**, 659–672 (1999).
- Landolt, A. UBVR photometric standard stars in the magnitude range 11.5–16.0 around the celestial equator. *Astron. J.* **104**, 340–371 (1992).
- Keller, H. U. et al. Comet P/Halley's nucleus and its activity. *Astron. Astrophys.* **187**, 807–823 (1987).
- Altenhoff, W. J. & Stumpff, P. Size estimate of 'asteroid' 2060 Chiron from 250GHz measurements. *Astron. Astrophys.* **293**, 41–42 (1995).
- Davies, J., Spencer, J., Sykes, M., Tholen, D. & Green, S. (5145) Pholus. *IAU Circular* 5698 (1993).
- Tholen, D. & Buie, M. in *Pluto and Charon* (eds Stern, S. A. & Tholen, D. J.) 193–220 (Univ. Arizona Press, Tucson, 1997).

Acknowledgements

This work was supported by the National Science Foundation. H.A. is supported by a James Clerk Maxwell Telescope fellowship. We thank the Institute for Astronomy Director for allocating target of opportunity JCMT time and D. Sanders for donating UH telescope time.

Correspondence and requests for materials should be addressed to D.J. (e-mail: jewitt@ifa.hawaii.edu).

An unusual phase transition to a second liquid vortex phase in the superconductor $\text{YBa}_2\text{Cu}_3\text{O}_7$

F. Bouquet^{*†}, C. Marcenat^{*}, E. Steep^{*}, R. Calemczuk[†], W. K. Kwok[‡], U. Welp[‡], G. W. Crabtree[‡], R. A. Fisher[†], N. E. Phillips[†] & A. Schilling[§]

^{*} Département de Recherche Fondamentale sur la Matière Condensée, Service de Physique, Magnétisme et Supraconductivité, CEA-Grenoble, 38054, Grenoble, France

[†] Department of Chemistry, University of California and Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

[‡] Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

[§] Physik-Institut der Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

A magnetic field penetrates a superconductor through an array of 'vortices', each of which carries one quantum of flux that is surrounded by a circulating supercurrent. In this vortex state, the resistivity is determined by the dynamical properties of the vortex 'matter'. For the high-temperature copper oxide superconductors (see ref.1 for a theoretical review), the vortex phase can be a 'solid', in which the vortices are pinned, but the solid can 'melt' into a 'liquid' phase, in which their mobility gives rise to a finite resistance. (This melting phenomenon is also believed to occur in conventional superconductors, but in an experimentally inaccessible part of the phase diagram².) For the case of $\text{YBa}_2\text{Cu}_3\text{O}_7$, there are indications of the existence of a critical point, at which the character of the melting changes^{3–10}. But neither the thermodynamic nature of the melting, nor the phase diagram in the vicinity of the critical point, has been well established. Here we report measurements of specific heat and magnetization that determine the phase diagram in this material to 26 T, well above the critical point. Our results reveal the presence of a reversible second-order transition above the critical point. An unusual feature of this transition—namely, that the high-temperature phase is the less symmetric in the sense of the

Landau theory¹¹—is in accord with theoretical predictions^{12–14} of a transition to a second vortex-liquid phase.

The melting of an ordinary solid is a thermodynamic first-order transition with discontinuities in physical properties, including the entropy (S). However, examples of second-order, continuous transitions with no discontinuity in S , are also common. For $\text{YBa}_2\text{Cu}_3\text{O}_7$ (YBCO), a hysteretic sharp drop in the resistivity³ gave evidence of first-order melting that terminated at a critical point. Thermodynamic first-order melting was suggested^{15,16} by a discontinuity in magnetization (M), and confirmed¹⁷ by measurements of both M and the specific heat (C). Later measurements^{5–10} of C showed a change in the melting at the critical point. That change is thought to be associated with a change in the nature of the solid^{18,19}, from an ordered lattice at low magnetic field (H) to a disordered glass-like solid at high H . The lattice–glass transition in the vortex matter is related to sample-dependent disorder in the underlying crystal structure and the vortex pinning that it produces. For $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (BSCCO), the melting of the vortex lattice was observed by local-field measurements²⁰. The lattice–glass transition is marked by a sharp feature²¹ in M at a field that extrapolates to the critical point, and by a change in the neutron diffraction pattern²². (Because the melting occurs at low H , the thermal effects are small and have not been detected.) For YBCO, which does not have the extreme anisotropy of BSCCO, the lattice–glass transition is not as sharp, and neither its thermodynamic nature nor that of the glass phase are well defined. (The terms 'glass phase' and 'lattice–glass transition' are used in the following only for convenience.) However, somewhat below the melting temperature, the 'fishtail', a broad second peak in the hysteretic magnetization loop (see below) that also seems to extrapolate to the critical point, has been identified²³ as a manifestation of the lattice–glass transition. Here we report measurements close to the melting line that reveal a deviation of the fishtail to lower H , giving insight into the nature of the melting and its relation to the lattice–glass transition.

The sample was a 1.5-mg, optimally doped (transition temperature $T_c \approx 92$ K) single crystal with small numbers of twin boundaries in two corners. Measurements of C were made with the a.c. technique²⁴. Heat is supplied at a frequency ω , typically ~ 1 Hz, to a sample loosely coupled to a controlled temperature block, inducing an alternating temperature response with a peak-to-peak amplitude T_{ac} . Measurements can be made while sweeping H or T , either up or

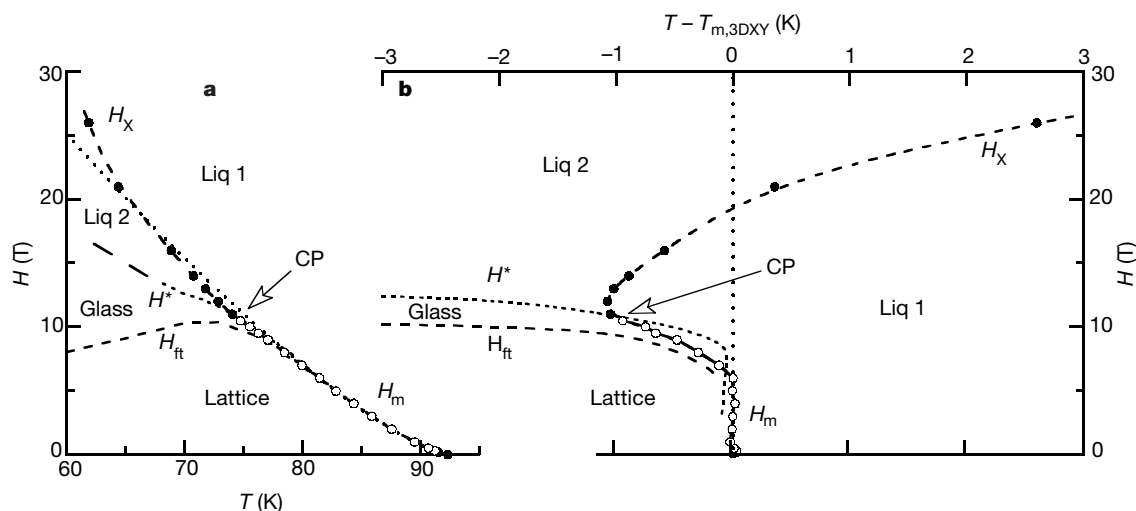


Figure 1 Phase diagram for the vortex state in YBCO. **a**, As H versus T ; **b**, on an expanded temperature scale, as deviations in T from a 3DXY scaling relation, the dotted line in both **a** and **b**, that represents the line of first-order melting for $H < 5$ T. H_m is the line of first-order melting, defined by discontinuities in entropy (at the open circles); H_x is the phase boundary between the two liquid phases, defined by discontinuities in specific heat (at the

solid circles); H_{ft} represents the maxima of the 'fishtail' in the magnetization loops, and indicates the region in which the lattice–glass transition occurs; H^* is a line below which irreversibility is important, defined by the points at which the magnetization loops close (below 68 K, where the line is dashed, the precision of the values of H is lower than at higher T). CP, critical point.

down. The sweep rate, ω and T_{ac} can all be varied to obtain information about time dependence and an estimate of the hysteresis at a phase transition (the difference in transition temperatures on cooling and warming)^{7,8}. Time-dependent irreversible effects do not appear in these measurements if the relaxation times are longer than ω^{-1} . The sensitivity of the technique as used in these measurements is typically ~ 1 in 10^4 for $T_{ac} \approx 40$ mK. Some magnetization measurements were made with a SQUID magnetometer but most were made with a torque magnetometer, which is better suited to the study of time-dependent effects and permits measurements to higher fields, where it is more sensitive.

The phase diagram, as defined by features in both C and M , is shown in Fig. 1. The features in $C(H)$ are shown in Fig. 2a, b and c, as $[C(H) - C(0)]/T$, to emphasize the H -dependent part, which is at most 0.8% of C . For low values of H there is a sharp 'peak' in C superimposed on a 'step' (ΔC), the height of which is the discontinuity in C between two phases. The area in C/T under the peak determines the discontinuity in entropy between the two phases (ΔS). The maxima of the peaks define the line of first-order melting, $T_m(H)$ or $H_m(T)$, which terminates at the critical point (CP), $H_{CP} = 10.5$ T, $T_{CP} = 74$ K, where $\Delta S \rightarrow 0$ (Fig. 1, and inset of Fig. 2a).

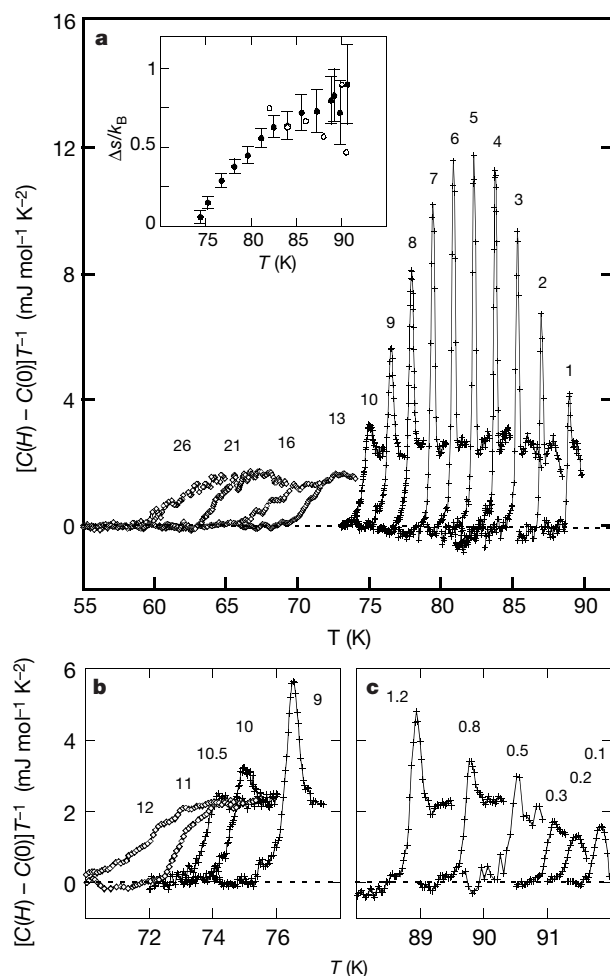


Figure 2 Discontinuities in specific heat and entropy at H_m and H_x . **a**, Specific-heat data for representative fields between 1 and 26 T; **b** and **c**, data for additional fields, on an expanded temperature scale, respectively in the vicinity of the critical point and for low fields. The data points have been shifted vertically to bring them into coincidence at low T . Inset in **a**, discontinuities in entropy at H_m , as measured directly (closed circles), and calculated from ΔM and the Clapeyron equation (open circles), per vortex per layer, in units of k_B .

Above the critical point, where there are no peaks, the steps persist, defining a line of second-order transitions, H_x . The width of the steps, which varies from less than 0.2 K at low H to 3 K at 26 T, is comparable to the widths of features in other properties of clean samples. Magnetization loops and their relation to features in C display the three types of behaviour shown in Fig. 3. The points at which the magnetization loops close define a line in the phase diagram, H^* , below which irreversibility is important. The second maximum in the amplitude of the magnetization loop (Fig. 3b), the fishtail, is represented by the H_{ft} line.

For $H \leq 5$ T, the features in C and M conform to all expectations for thermodynamic first-order melting of a disorder-free lattice: the melting follows the 3DXY scaling relation, $T_m(H) = T_{m,3DXY} = T_c[1 - (H/H_0)^{3/4}]$ (Fig. 1), with $T_c = 92.3$ K and $H_0 = 100.3$ T. It occurs entirely in the reversible region (Figs 1 and 3a), permitting the measurement of ΔM . As shown in the inset in Fig. 2a, where they are compared as the entropy change per vortex per layer (Δs), the values of ΔS measured directly and those calculated from ΔM and the Clapeyron equation, $dH_m/dT = -\Delta S/\Delta M$, are in good agreement, confirming thermodynamic equilibrium. The measurements of C show the presence of hysteresis⁷⁻⁹, which often accompanies a first-order transition but is not expected for a second-order transition. In this region, the width of the melting transition, its hysteresis, and Δs , are all constant.

As H increases from 5 T to H_{CP} , the parameters that characterize

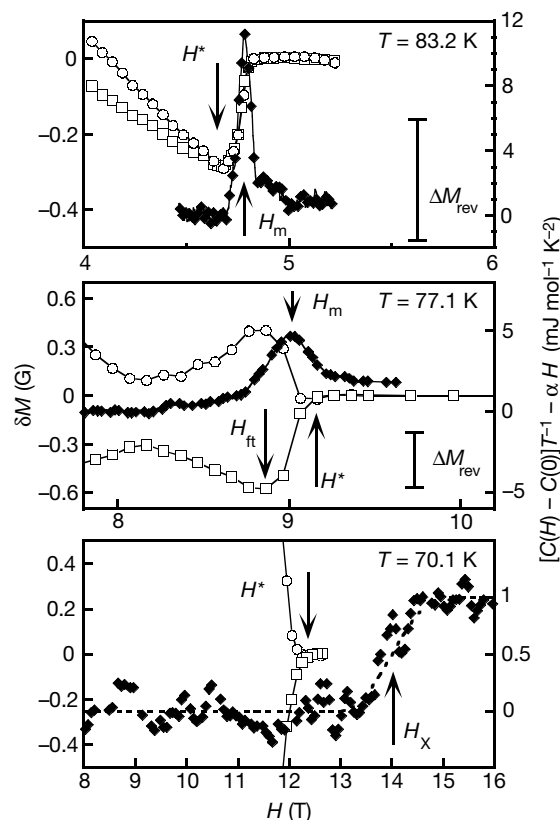


Figure 3 Magnetization and specific-heat data that are characteristic of three regions of the phase diagram. (Open squares and open circles show data for increasing and decreasing field, respectively.) In each case a (different) smooth background has been subtracted to put the data on a more convenient scale, and the residual, δM , is plotted. H^* and H_{ft} mark the closing of the magnetization loop and the maximum of the fishtail, respectively. Specific-heat data (solid diamonds), taken in field sweeps at the same temperatures, are shown for comparison. A term proportional to H , which fits the low- H data, has been subtracted. ΔM_{rev} represents the magnitude of the discontinuity in M expected from the discontinuity in entropy and the Clapeyron equation.

the melting change gradually: the melting temperature falls below $T_{m,3DXY}$, the value extrapolated from $H \leq 5$ T (Fig. 1); ΔS decreases, to zero at the critical point; and both the hysteresis and the width of the melting transition (compare Fig. 3a and 3b) increase. These changes are accompanied by a change in the relation between the thermal signature of the melting and the onset of irreversibility (Figs 1 and 3). Above 5 T, the onset of irreversibility moves into the region of melting and the amplitude of the magnetization loop increases substantially. There is no observable ΔM , and the thermodynamic nature of the transition cannot be tested by the Clapeyron equation. Nevertheless, the values of ΔS obtained from the C measurements are the same for H and T sweeps and independent of their direction, indicating that they correspond to equilibrium.

The change in the character of the melting is accompanied by a change in its relation to the lattice–glass transition, which is represented by the fishtail. The high sensitivity of the measurements made possible the observation of the fishtail to 81 K and 5 T where it merges with the melting line, well beyond the region in which it had been seen previously. (At 81 K its amplitude is an order of magnitude smaller than at 77 K, for which it is shown in Fig. 3b, and more than three orders of magnitude smaller than at 70.1 K, where its maximum is far off-scale in Fig. 3c.) At low T the H_{fi} line seems to extrapolate to the critical point (Fig. 1 and ref. 23), but it develops a negative curvature as it approaches the melting line, producing an extended range of H , precisely that in which the character of the melting changes, where it is in close proximity to the melting line. The relation between the melting and H_{fi} lines in this region of the phase diagram suggests that the changes in the melting parameters with increasing H , in particular the decrease in ΔS , can be understood in terms of melting from an increasingly disordered solid.

The onset of irreversibility in M and the onset of finite resistivity are both intimately related to the melting of the vortex solid. The H^* line, with its change of slope at the critical point (Fig. 1a), mirrors precisely the melting lines for the glass and lattice phases in the vicinity of the critical point deduced from resistivity measurements³ on a sample with a similar H_{CP} ; this behaviour supports the conclusion that for $H \approx H_{CP}$ the ‘melting’ of the glass occurs in the vicinity of the onset of irreversibility, near H^* . Measurements of C by a different technique, similar to that in ref. 17, show time-dependent, irreversible features in this region, but, as expected, they are not seen in the a.c. measurements reported here. As a consequence of the high sensitivity of the M measurements (~ 2 mG) and the steepness with which the hysteresis loop closes (Fig. 3c), H^* defines a relatively sharp boundary between this region in which irreversibility is important, and a higher- T region in which it is not.

The features that are seen in the a.c. measurements of C above H_{CP} mark a transition that is qualitatively different from the melting of the vortex lattice below H_{CP} . Below H_{CP} the peak/steps in C (at H_m) show hysteresis, which increases as the critical point is approached; they occur close to H^* , with H_m and H^* differing by at most a small fraction of a kelvin (Fig. 1b). Above H_{CP} the steps in C (at H_X) occur entirely within the reversible region of the M measurements, well separated from H^* (Fig. 3c); they do not show hysteresis, and the values of ΔC are the same for H and T sweeps, in either direction. At the critical point, where H^* bends to lower H , H_X diverges from it, to higher H (Fig. 1a and b). At the H_X line the values of ΔC are independent of ω (in the range explored, 0.5–10 Hz), and the relaxation times in measurements of M are not detectable (they become detectable, at ~ 10 s, only several kelvins below the H_X line). In earlier reports^{5–9}, steps in C above the critical point were interpreted as manifestations of a second-order glass–liquid transition, but that is inconsistent with the absence of significant time dependence in the properties measured at the H_X line (see, for example, refs 18 and 25). The combination of measurements of C and M and their time dependences reported here show that H_X defines a line of reversible, thermodynamic, second-order transitions that requires a new interpretation¹⁰. In the Landau

theory of second-order transitions¹¹ the less symmetric phase, with the non-zero order parameter, has the higher specific heat. Within the confines of the Landau theory, the sign of ΔC requires that the high- T phase be the less symmetric, which is very unusual, and imposes a significant constraint on the interpretation of the transition. For example, while in general the change in symmetry accompanying the melting of a glass to a liquid is not well defined, at least for one particular model¹⁸ the liquid is the more symmetric phase.

This unusual feature of the transition at H_X is predicted by recent theoretical work: both analytical calculations based on general duality arguments^{12,13} and computer simulations¹⁴ reveal a second-order transition with the observed sign of ΔC , from a low- T vortex liquid with line tension, to a high- T vortex liquid without line tension. The order parameter, which is related to the connectivity of the vortex system, is zero in the low- T liquid and non-zero in the high- T liquid where there are unbound vortex loops threading the whole system. The mechanism, a ‘vortex loop blow-out’, has also been invoked^{26,27} in connection with the properties of superfluid ^4He , and the underlying theory has implications^{13,14,26} for other physical phenomena. In the computer simulations¹⁴ the liquid–liquid and melting transition lines intersect at a non-zero characteristic value of H ; in the analytical calculations¹³ that occurs at $H = 0$, but the two transitions may coincide if the melting is ‘strongly’ first-order. Because effects of disorder are not included in either work, any detailed comparison with experiment must be considered tentative. However, a possible interpretation of the experimental results, consistent with the theory in its present form, is that the two transitions are coincident for $H < H_{CP}$ but separate at H_{CP} as a consequence of the ‘weakening’ of the melting transition. This would account for the intersection of the H_X and melting lines at the critical point, and the similarity of ΔC above and below the critical point. □

Received 3 October 2000; accepted 21 March 2001.

- Blatter, G., Feigel'man, M. V., Geshkenbein, V. B., Larkin, A. I. & Vinokur, V. M. Vortices in high-temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1388 (1994).
- Eilenberger, G. Thermodynamic fluctuations of the order parameter in type-II superconductors near the upper critical field H_{c2} . *Phys. Rev.* **164**, 628–635 (1967).
- Safar, H. *et al.* Experimental evidence for a multicritical point in the magnetic phase diagram for the mixed state of clean, untwinned $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. Lett.* **70**, 3800–3803 (1993).
- Nishizaki, T., Naito, T. & Kobayashi, N. Anomalous magnetization and field-driven disordering transition of a vortex lattice in untwinned $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys. Rev. B* **58**, 11169–11172 (1998).
- Roulin, M., Junod, A., Erb, A. & Walker, E. Calorimetric transitions on the melting line of the vortex system as a function of oxygen deficiency in high purity $\text{YBa}_2\text{Cu}_3\text{O}_x$. *Phys. Rev. Lett.* **80**, 1722–1725 (1998).
- Roulin, M. *et al.* in *Physics and Materials Science of Vortex State, Flux Pinning and Dynamics* (eds Kossowsky, R. *et al.*) 489–504 (Kluwer Academic, Dordrecht, 1999).
- Billon, B. *Étude Microcalorimétrique et Magnétique du Diagramme de Phase des Composés $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ et $\text{HgBa}_2\text{CuO}_{4+x}$* . Thesis, Univ. Grenoble 1 (1998).
- Bouquet, F. *Étude thermodynamique de l'état mixte dans $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$* . Thesis, Univ. Grenoble 1 (1998).
- Bouquet, F. *et al.* in *Physics and Materials Science of Vortex State, Flux Pinning and Dynamics* (eds Kossowsky, R. *et al.*) 743–758 (Kluwer Academic, Dordrecht, 1999).
- Marcat, C. *et al.* The vortex state of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$: phase diagram; calorimetric and magnetic properties. *Physica C* **341–348**, 949–952 (2000).
- Landau, L. D. & Lifshitz, E. M. in *Course of Theoretical Physics* 2nd edn Vol. 5, *Statistical Physics Ch. VIII* 257 (Pergamon, Oxford, 1969).
- Tešanović, Z. Low-magnetic-field critical behavior in strongly type-II superconductors. *Phys. Rev. B* **51**, 16204–16210 (1995).
- Tešanović, Z. Extreme type-II superconductors in a magnetic field: A theory of critical fluctuations. *Phys. Rev. B* **59**, 6449–6474 (1999).
- Nguyen, A. K. & Sudbø, A. Topological phase fluctuations, amplitude fluctuations, and criticality in extreme type-II superconductors. *Phys. Rev. B* **60**, 15307–15331 (1999).
- Liang, R., Bonn, D. A. & Hardy, W. N. Discontinuities of reversible magnetization in untwinned YBCO single crystals at the first order vortex melting transition. *Phys. Rev. Lett.* **76**, 835–838 (1996).
- Fendrich, J. A. *et al.* Static and dynamic vortex phases in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Phys. Rev. Lett.* **77**, 2073–2076 (1996).
- Schilling, A. *et al.* Calorimetric measurement of the latent heat of vortex-lattice melting in untwinned $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Nature* **382**, 791–793 (1996).
- Fisher, D. S., Fisher, M. P. A. & Huse, D. A. Thermal fluctuations, quenched disorder, phase transitions and transport in type-II superconductors. *Phys. Rev. B* **43**, 130–159 (1991).
- Giamarchi, T. & Le Doussal, P. Phase diagrams of flux lattices with disorder. *Phys. Rev. B* **55**, 6577–6583 (1997).
- Zeldov, E. *et al.* Thermodynamic observation of first-order vortex-lattice melting transition in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$. *Nature* **375**, 373–376 (1995).

21. Khaykovich, B. *et al.* Vortex-lattice phase transitions in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ crystals with different oxygen stoichiometry. *Phys. Rev. Lett.* **76**, 2555–2558 (1996).
22. Cubitt, R. *et al.* Direct observation of magnetic flux lattice melting and decomposition in the high- T_c superconductor $\text{Bi}_{1.5}\text{Sr}_{1.95}\text{CaCu}_2\text{O}_{8+x}$. *Nature* **365**, 407–411 (1993).
23. Deligiannis, K. *et al.* New features in the vortex phase diagram of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. *Phys. Rev. Lett.* **79**, 2121–2124 (1997).
24. Sullivan, P. F. & Seidel, G. Steady-state, ac-temperature calorimetry. *Phys. Rev.* **173**, 679–685 (1968).
25. Reichhardt, C., van Otterlo, A. & Zimányi, G. T. Vortices freeze like window glass: the vortex molasses scenario. *Phys. Rev. Lett.* **84**, 1994–1997 (2000).
26. Williams, G. A. Vortex-loop phase transitions in liquid helium, cosmic strings, and high- T_c superconductors. *Phys. Rev. Lett.* **82**, 1201–1204 (1999).
27. Onsager, L. Statistical hydrodynamics. *Nuovo Cimento Suppl.* **6**, 279–287 (1949).

Acknowledgements

The measurements above 18 T were performed in the Grenoble High Magnetic Field Laboratory. Work at Argonne and Berkeley was supported by the US DOE.

Correspondence and requests for materials should be addressed to F.B.
(e-mail: Frederic_Bouquet@Yahoo.Com).immuno

‘Inverse’ melting of a vortex lattice

Nurit Avraham*, Boris Khaykovich†, Yuri Myasoedov*,
Michael Rappaport*, Hadas Shtrikman*, Dima E. Feldman*‡,
Tsuyoshi Tamegai§, Peter H. Kes||, Ming Li||, Marcin Konczykowski¶,
Kees van der Beek¶ & Eli Zeldov*

* Department of Condensed Matter Physics, The Weizmann Institute of Science, Rehovot 76100, Israel

† Landau Institute for Theoretical Physics, 142432 Chernogolovka, Moscow region, Russia

§ Department of Applied Physics, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, and CREST, Japan Science and Technology Corporation (JST), Japan

|| Kamerlingh Onnes Laboratory, Leiden University, PO Box 9504, 2300 RA Leiden, The Netherlands

¶ Laboratoire des Solides Irradiés, CNRS, UMR 7642 and CEA/DSM/DRECAM, Ecole Polytechnique, 91128 Palaiseau, France

Inverse melting is the process in which a crystal reversibly transforms into a liquid or amorphous phase when its temperature is decreased. Such a process is considered to be very rare¹, and the search for it is often hampered by the formation of non-equilibrium states or intermediate phases². Here we report the discovery of first-order inverse melting of the lattice formed by magnetic flux lines in a high-temperature superconductor. At low temperatures, disorder in the material pins the vortices, preventing the observation of their equilibrium properties and therefore the determination of whether a phase transition occurs. But by using a technique³ to ‘dither’ the vortices, we were able to equilibrate the lattice, which enabled us to obtain direct thermodynamic evidence of inverse melting of the ordered lattice into a disordered vortex phase as the temperature is decreased. The ordered lattice has larger entropy than the low-temperature disordered phase. The mechanism of the first-order phase transition changes gradually from thermally induced melting at high temperatures to a disorder-induced transition at low temperatures.

Local magnetization measurements, using microscopic Hall sensors⁴, were performed on a number of high-quality optimally doped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (BSCCO) crystals, grown in two laboratories^{5,6}. A d.c. magnetic field H_a was applied parallel to the crystalline c axis, while an a.c. transverse field $H_{ac\perp}$ was applied along the a – b planes (see Fig. 1 legend for details).

In $\text{YBa}_2\text{Cu}_3\text{O}_7$ at temperatures close to T_c , ‘vortex dithering’ by a transverse a.c. field reduces the irreversible magnetization caused by vortex pinning³. We find that in BSCCO crystals this method can fully suppress the magnetic hysteresis even at low temperatures, down to about 30 K, as shown in Fig. 1a. The Abrikosov vortices in BSCCO can be regarded as a stack of Josephson-coupled pancake vortices⁷ in the individual CuO_2 planes. The in-plane field $H_{ac\perp}$ readily penetrates through the sample⁸ in the form of Josephson vortices residing between the CuO_2 planes⁹. The main effect, which is of interest here, is that a pancake vortex, located in a CuO_2 plane immediately above a Josephson vortex, is displaced a small distance

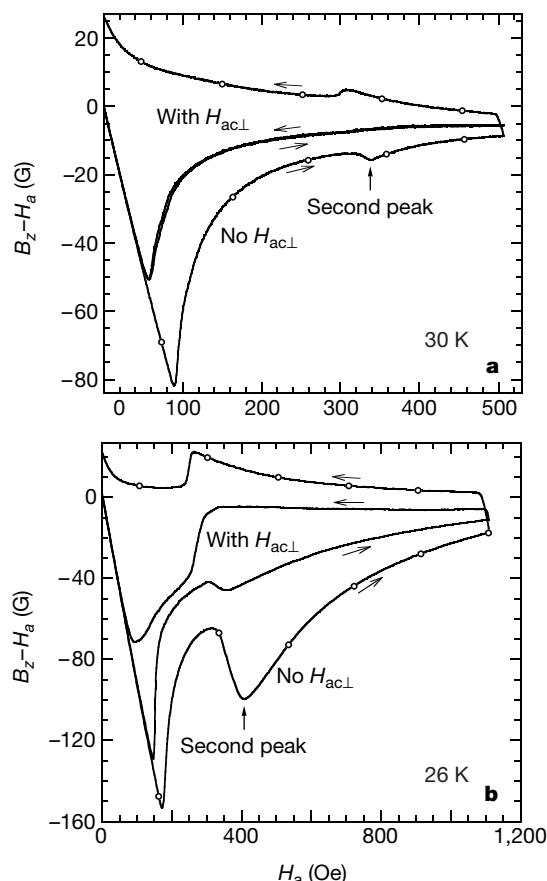


Figure 1 Local magnetization loops in BSCCO crystals with and without (open circles) ‘vortex dithering’. A number of optimally doped BSCCO crystals with $T_c \approx 90$ K were investigated. We present here results on three crystals: A ($160 \times 600 \times 25 \mu\text{m}^3$), B ($170 \times 600 \times 20 \mu\text{m}^3$) and C ($190 \times 900 \times 30 \mu\text{m}^3$). The samples were mounted on an array of $10 \times 10 \mu\text{m}^2$ GaAs/AlGaAs Hall sensors. A d.c. magnetic field H_a was applied perpendicular to the surface of the sensors (parallel to the crystalline c axis). A transverse a.c. field $H_{ac\perp}$ (1 kHz with amplitude of up to 80 Oe), applied parallel to the ab planes of the crystals, equilibrates the vortices by ‘dithering’ and suppresses the magnetic hysteresis. The Hall sensors are insensitive to the transverse field and measure only the d.c. c -axis component B_z of the local induction. **a**, Lowest temperature, 30 K, at which fully reversible magnetization in crystal A is attained for our maximum $H_{ac\perp}$ of ~ 80 Oe. The d.c. magnetization loop displays the second magnetization peak above 300 Oe, which is due to the enhanced vortex pinning in the disordered phase. Application of $H_{ac\perp}$ results in a reversible magnetization curve that reveals a small step in magnetization (barely seen on this scale) at the original location of the second magnetization peak. **b**, Example of the local magnetization loop at a lower temperature (26 K) in crystal B. The d.c. loop shows large hysteresis and a pronounced second magnetization peak. The applied $H_{ac\perp}$ is not sufficient to obtain reversible magnetization. The ‘dithering’ efficiency of $H_{ac\perp}$ is significantly different in the two vortex phases: For the same width of the original d.c. loop, much stronger suppression of the hysteresis is observed in the ordered phase than in the disordered phase above the second magnetization peak.

† Present address: Department of Physics and CMSE, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

along the direction of the Josephson vortices relative to the neighbouring pancake vortex residing one CuO_2 layer underneath⁹. For small values of $H_{ac\perp}$, the probability of such a close proximity between a pancake vortex and a Josephson vortex is low. Thus, in our sample geometry, an average pancake vortex experiences only a few such periodic 'intersections' during one a.c. cycle; the duration of each intersection is less than 1% of the a.c. period. $H_{ac\perp}$ therefore induces a weak local a.c. agitation of pancake vortices, which assists thermal activation in relaxing the irreversible magnetization and in approaching thermal equilibrium.

In addition to bulk pinning, a substantial part of the magnetic hysteresis in BSCCO crystals is caused by surface¹⁰ and geometrical barriers¹¹. $H_{ac\perp}$ also efficiently suppresses this source of hysteresis. The Josephson vortices, which periodically cross the sample edges, instantaneously 'slice' the Abrikosov vortices into short segments or individual pancakes, thus apparently facilitating¹⁰ vortex penetration through the surface and geometrical barriers. At lower temperatures the equilibration of vortices by $H_{ac\perp}$ becomes progressively more difficult, as shown in Fig. 1b at $T = 26$ K, where our maximum $H_{ac\perp}$ is not sufficient to eliminate the hysteresis completely.

We now focus on the main topic of vortex-matter phase transitions. At elevated temperatures the vortex lattice undergoes a first-order melting transition^{4,8,12,13}. The melting line $B_m(T)$ apparently terminates at some critical point T_{CP} , below which a 'second magnetization peak' line $B_{sp}(T)$ emerges^{14–17}. In BSCCO crystals the $B_{sp}(T)$ line is rather horizontal on the B – T phase diagram. On crossing this line by a field sweep, a pronounced 'second peak' is observed in magnetization loops¹⁴ (Fig. 1), which reflects a transformation of a quasi-ordered lattice into a disordered amorphous state with enhanced vortex pinning.

The first-order nature of the melting transition $B_m(T)$ is manifested by a small step in equilibrium magnetization, as described in Fig. 2. At high temperatures, $H_{ac\perp}$ has no observable effect on this step. As the temperature is decreased along $B_m(T)$ towards T_{CP} , magnetic hysteresis starts to set in gradually, as seen for example in Fig. 2a. Here, at $T > T_{CP}$, the magnetization step is still clearly resolved, but it is partly suppressed by the hysteresis in the vortex solid phase below the transition. Application of $H_{ac\perp}$ fully removes the hysteresis and enhances the magnetization step. $H_{ac\perp}$ also results in a small broadening of the transition. We attribute this broadening to the fact that a transverse field slightly decreases⁹ the melting field B_m . The small variation of B_m during the a.c. cycle of $H_{ac\perp}$ should therefore cause some smearing of the magnetization step.

Figure 2b shows the behaviour at $T = 38$ K, just below T_{CP} , where the first-order transition was believed to be absent. At this temperature, significant hysteresis is observed and the second magnetization peak starts to develop. Application of $H_{ac\perp}$ eliminates the hysteresis and instead a step in reversible magnetization emerges, which reveals the existence of a first-order transition. A similar phenomenon is found at lower temperatures, as shown in Fig. 2c, where a clear second magnetization peak turns into a step in magnetization upon application of $H_{ac\perp}$ (see also Fig. 1a). This finding leads to several important conclusions, including the following: (1) the first-order transition does not terminate at T_{CP} , (2) the vortex matter displays inverse melting, (3) the underlying mechanism of the first-order transition is different above and below T_{CP} , and (4) the disorder-driven transition at the second magnetization peak is of first order.

(1) The first conclusion is that the previously reported termination of the first-order transition⁴ at T_{CP} does not reflect a real critical point, but rather an experimental limitation. Below T_{CP} the dynamics becomes too slow to achieve vortex equilibration on typical experimental timescales, thus preventing observation of the equilibrium magnetization step. $H_{ac\perp}$ accelerates the equilibration process³ and thus facilitates the detection of the magnetization step down to significantly lower temperatures.

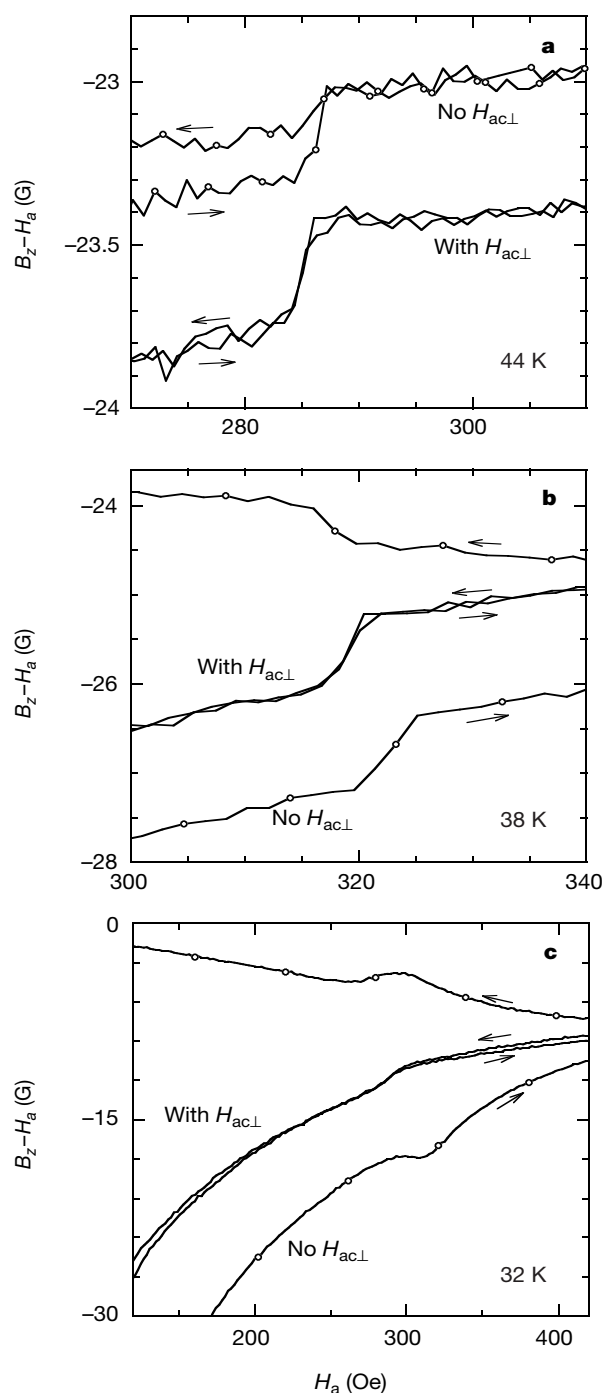


Figure 2 Reversible magnetization steps revealed in BSCCO crystals by 'vortex dithering' at various temperatures. At high temperatures the first-order magnetization step occurs in the region of reversible magnetization. In this case $H_{ac\perp}$ has no effect on the magnetization step up to our maximum transverse field of 80 Oe. In addition, there are no heating effects, as confirmed by the absence of a temperature shift in $B_m(T)$. **a**, At intermediate temperatures, such as $T = 44$ K $> T_{CP} \approx 40$ K (crystal C), hysteretic magnetization is observed below the transition (open circles). Application of $H_{ac\perp}$ removes the residual hysteresis and enhances the magnetization step. Here the curve with $H_{ac\perp}$ was displaced downwards for clarity. **b**, At temperatures slightly below T_{CP} , significant hysteresis develops in the d.c. magnetization loops and the beginning of the second-magnetization-peak behaviour is observed in the form of a pronounced change in the width of the loop. $H_{ac\perp}$ results in reversible magnetization and reveals a clear magnetization step that appears instead of the second magnetization peak (crystal C, $T = 38$ K). **c**, At still lower temperatures a fully developed second magnetization peak is observed in the d.c. magnetization loop which is replaced by a step in the reversible magnetization in the presence of $H_{ac\perp}$ (crystal B, $T = 32$ K).

(2) Observation of the magnetization step below T_{CP} is not simply an extension of the range of the first-order transition, but rather provides qualitatively new information. Figure 3 shows the location of the transition line on the B - T phase diagram. Without $H_{ac\perp}$, the first-order transition is observed only at temperatures above T_{CP} , where $B_m(T)$ has a negative slope dB_m/dT . We note that $B_m(T)$ becomes horizontal⁴ on approaching T_{CP} . By applying $H_{ac\perp}$ we find that the first-order transition extends to significantly lower temperatures and shows an inverted behaviour of a positive slope below T_{CP} . Hence, in this region the vortex matter displays the very rare phenomenon of inverse melting in which an ordered lattice (Bragg glass^{18,19}) reversibly transforms into a liquid or amorphous phase upon cooling.

(3) In a thermally driven vortex-melting transition, the liquid phase has to be on the high-temperature side of the transition because thermal fluctuations always increase with temperature. The observed inverse melting behaviour therefore implies that the first-order transition below T_{CP} must have a different underlying nature, which is disorder-driven rather than thermally driven. The main conceptual difference is that in thermal melting the decrease in elastic energy of the lattice at the transition is compensated for by a gain in entropy of the liquid. In a disorder-driven transition, in contrast, the loss of elastic energy is balanced by a gain in pinning energy, because the 'wiggling' entangled vortices adapt more efficiently to the pinning landscape induced by the quenched material disorder²⁰. At low fields, the elastic energy of the lattice is higher than the pinning energy, giving rise to an ordered state¹⁸⁻²⁰. With increasing field the elastic energy decreases relative to pinning, leading to a structural transformation into a disordered phase when the two energies become comparable. Our finding shows that this non-thermal transition is of first order. At low temperatures the shape of the disorder-driven transition line should be nearly temperature independent or may show downward curvature due to the temperature dependence of the microscopic parameters²¹. At intermediate temperatures, however, even though the transition remains disorder-driven, thermal smearing of the pinning potential progressively reduces the pinning energy, leading to an upturn in the transition line²⁰. This effect of pinning reduction is apparently at

the origin of the observed inverse melting behaviour. An upturn in the shape of the second magnetization peak line has been noted previously^{5,14}; however, the results presented here are, to our knowledge, the first evidence of a thermodynamic inverse melting transition of the vortex lattice.

(4) The extended $B_m(T)$ transition line coincides with the location of the $B_{sp}(T)$ line at lower temperatures, as seen in the inset to Fig. 3, which demonstrates that the two phenomena have a common origin. The second magnetization peak is the dynamic characteristic of the transition, reflecting the enhanced vortex pinning and the viscous nature of the disordered amorphous phase, whereas the reversible magnetization step is its thermodynamic signature. The revealed first-order nature of the second peak is consistent with recent conclusions based on Josephson plasma resonance studies²² and transient measurements^{23,24}, as well as several numerical simulations²⁵⁻²⁷. Thus, the breakdown of the Bragg glass is apparently always a first-order transition that occurs through a thermal and/or a disorder-driven mechanism. Whereas thermally driven melting is possibly restricted to high-temperature superconductors, the disorder-driven transition should be a more general phenomenon and is apparently at the heart of the ubiquitous peak effect in low- T_c superconductors²⁸.

Figure 4 shows the magnetization step ΔB and the corresponding entropy change $\Delta S = -(dH_m/dT)\Delta B/4\pi$ in the vicinity of T_{CP} . ΔB is approximately constant in this region. The entropy change ΔS vanishes at T_{CP} because $dH_m/dT = 0$, and becomes negative below T_{CP} . A negative ΔS means that, paradoxically, the ordered lattice has larger entropy than the disordered phase. Because the ordered lattice has no dislocations and is structurally more ordered, the extra entropy must arise from additional degrees of freedom. In the amorphous phase the lattice structure is broken and the vortices wander out of their unit cells and become entangled, yet their thermal fluctuations on short timescales are small due to enhanced pinning. In contrast, in the ordered lattice there is no entanglement and no large-scale vortex wandering; however, thermal fluctuations within the unit cell are larger, apparently resulting in larger entropy.

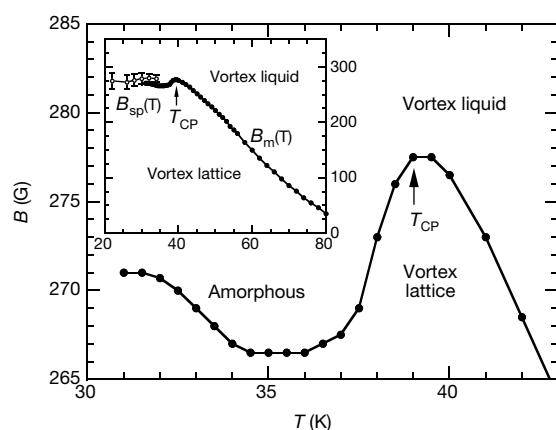


Figure 3 The first-order transition line and the inverse melting obtained with 'vortex dithering'. Inset, the first-order $B_m(T)$ line (filled circles) along with the second magnetization peak line $B_{sp}(T)$ (open circles) over a wide temperature range in BSCCO crystal B. $B_{sp}(T)$ was measured without $H_{ac\perp}$ and the error bars reflect the uncertainty in the determination of the location of the B_{sp} transition owing to its larger width in comparison with the width of the first-order magnetization step. Main panel, an expanded view of the first-order transition line in the vicinity of T_{CP} . The first-order transition was observed at all temperatures at which fully reversible magnetization was achieved by 'vortex dithering'. At $T < 31$ K no reversible magnetization could be attained in this crystal with our maximum $H_{ac\perp}$ and hence no magnetization step could be revealed.

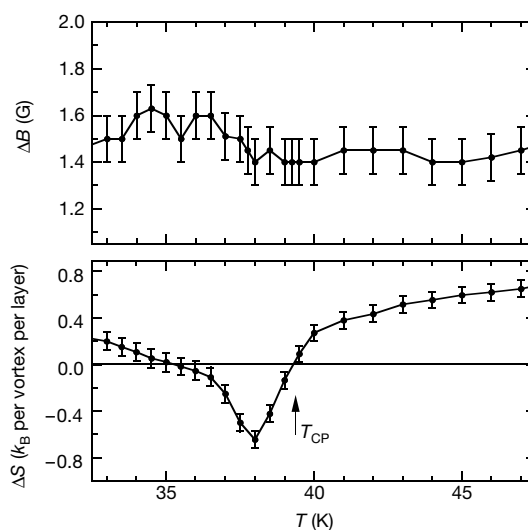


Figure 4 Magnetization step and the negative latent heat. The top panel is the height of the magnetization step ΔB in BSCCO crystal B in the vicinity of T_{CP} obtained in the presence of $H_{ac\perp}$. The bottom panel shows the corresponding calculated entropy change ΔS (or the latent heat $T\Delta S$). The negative ΔS below T_{CP} is the region of the inverse melting transition of the vortex lattice. In this region, on increasing field, the melting of the lattice is accompanied by a negative latent heat. However, on crossing this transition line by increasing the temperature, the opposite process of crystallization of the amorphous phase occurs, which is accompanied by a positive latent heat.

$\Delta S < 0$ also means that the latent heat $T_m \Delta S$ is negative, and hence the sample releases heat upon melting on increasing the field, in contrast to the usual melting in which the sample absorbs heat. Note, however, that by crossing the transition lines by increasing the temperature, the latent heat is always positive, as dictated by thermodynamics, and the high-temperature phase always has larger entropy. The observation of a rather constant ΔB in Fig. 4 resolves another previously reported⁴ inconsistency in the behaviour near T_{CP} : at a true critical point ΔB is expected to vanish continuously, whereas experimentally a rather constant ΔB was found to disappear abruptly at T_{CP} . The consistent values of ΔB above and below T_{CP} clearly indicate a smooth continuation of the first-order transition and the absence of a real critical point. However, we emphasize that the vortex system may not be in full thermal equilibrium in the presence of H_{ac} agitation. Therefore the relatively large values of ΔB in Fig. 4, which are found to be sample dependent, should be investigated further.

Finally, an important open question is whether the two disordered states of the vortex matter, liquid and amorphous, represent different thermodynamic phases or just a continuous slowing down of vortex dynamics. Recent numerical calculations led to contradictory conclusions^{25–27,29}. We do not observe any sharp features along the depinning line T_d , which usually extends upwards³⁰ from T_{CP} and is believed to separate the two phases. A strongly first-order T_d transition would require a sharp change in slope of $B_m(T)$ at T_{CP} , which we do not observe. Therefore, the T_d line in BSCCO is unlikely to be a first-order transition; it could be either a continuous transition or a dynamic crossover. Similarly, in $YBa_2Cu_3O_7$ the structure of the phase diagram is controversial, and it is as yet unclear whether the second magnetization peak line merges with the melting line at the critical point¹⁶ or at a lower field. □

Received 4 January; accepted 21 March 2001.

- Greer, A. L. Too hot to melt. *Nature* **404**, 134–135 (2000).
- Rastogi, S., Höhne, G. W. H. & Keller, A. Unusual pressure-induced phase behaviour in crystalline poly(4-methylpentene-1): calorimetric and spectroscopic results and further implications. *Macromolecules* **32**, 8897–8909 (1999).
- Willemin, M. *et al.* First-order vortex-lattice melting transition in $YBa_2Cu_3O_7$ near the critical temperature detected by magnetic torque. *Phys. Rev. Lett.* **81**, 4236–4239 (1998).
- Zeldov, E. *et al.* Thermodynamic observation of first-order vortex-lattice melting transition. *Nature* **375**, 373–376 (1995).
- Ooi, S., Shibauchi, T. & Tamegai, T. Evolution of vortex phase diagram with oxygen-doping in $Bi_2Sr_2CaCu_2O_{8+\delta}$ single crystals. *Physica C* **302**, 339–345 (1998).
- T. W. Li *et al.* Growth of $Bi_2Sr_2CaCu_2O_{8+\delta}$ single-crystals at different oxygen ambient pressures. *J. Cryst. Growth* **135**, 481–486 (1994).
- Blatter, G. *et al.* Vortices in high-temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1388 (1994).
- Pastoriza, H., Goffman, M. F., Arribere, A. & de la Cruz, F. First order phase transition at the irreversibility line of $Bi_2Sr_2CaCu_2O_8$. *Phys. Rev. Lett.* **72**, 2951–2954 (1994).
- Koshelev, A. E. Crossing lattices, vortex chains, and angular dependence of melting line in layered superconductors. *Phys. Rev. Lett.* **83**, 187–190 (1999).
- Burlachkov, L., Koshelev, A. E. & Vinokur, V. M. Transport properties of high-temperature superconductors: surface vs bulk effect. *Phys. Rev. B* **54** 6750–6757 (1996).
- Zeldov, E. *et al.* Geometrical barriers in high-temperature superconductors. *Phys. Rev. Lett.* **73**, 1428–1431 (1994).
- Safar, H. *et al.* Experimental evidence for a first-order vortex-lattice-melting transition in untwinned single crystal $YBa_2Cu_3O_7$. *Phys. Rev. Lett.* **69**, 824–827 (1992).
- Kwok, W. K. *et al.* Vortex lattice melting in untwinned and twinned single-crystals of $YBa_2Cu_3O_{7-\delta}$. *Phys. Rev. Lett.* **69**, 3370–3373 (1992).
- Khaykovich, B. *et al.* Vortex lattice phase transitions in $Bi_2Sr_2CaCu_2O_8$ crystals with different oxygen stoichiometry. *Phys. Rev. Lett.* **76**, 2555–2558 (1996).
- Chikumoto, N., Konczykowski, M., Motohira, N. & Malozemoff, A. P. Flux-creep crossover and relaxation over surface barriers in $Bi_2Sr_2CaCu_2O_8$ crystals. *Phys. Rev. Lett.* **69**, 1260–1263 (1992).
- Deligiannis, K. *et al.* New features in the vortex phase diagram of $YBa_2Cu_3O_7$. *Phys. Rev. Lett.* **79**, 2121–2124 (1997).
- Cubitt, R. *et al.* Direct observation of magnetic flux lattice melting and decomposition in the high- T_c superconductor $Bi_2Sr_2CaCu_2O_8$. *Nature* **365**, 407–411 (1993).
- Giamarchi, T. & Le Doussal, P. Elastic theory of pinned flux lattice. *Phys. Rev. Lett.* **72**, 1530–1533 (1994).
- Nattermann, T. & Scheidl, S. Vortex-glass phases in type-II superconductors. *Adv. Phys.* **49**, 607–704 (2000).
- Ertas, D. & Nelson, D. R. Irreversibility, entanglement and thermal melting in superconducting vortex crystals with point impurities. *Physica C* **272**, 79–86 (1996).
- Giller, D. *et al.* Disorder-induced transition to entangled vortex-solid in Nd-Ce-Cu-O crystal. *Phys. Rev. Lett.* **79**, 2542–2545 (1997).
- Gaifullin, M. B. *et al.* Abrupt change of Josephson plasma frequency at the phase boundary of the

- Bragg glass in $Bi_2Sr_2CaCu_2O_{8+\delta}$. *Phys. Rev. Lett.* **84**, 2945–2948 (2000).
- van der Beek, C. J., Colson, S., Indenbom, M. V. & Konczykowski, M. Supercooling of the disordered vortex lattice in $Bi_2Sr_2CaCu_2O_{8+\delta}$. *Phys. Rev. Lett.* **84**, 4196–4199 (2000).
- Giller, D., Shaulov, A., Tamegai, T. & Yeshurun, Y. Transient vortex states in $Bi_2Sr_2CaCu_2O_{8+\delta}$ crystals. *Phys. Rev. Lett.* **84**, 3698–3701 (2000).
- Kierfeld, J. & Vinokur, V. Dislocations and the critical endpoint of the melting line of vortex line lattices. *Phys. Rev. B* **61**, R14928–R14931 (2000).
- Nonomura, Y. & Hu, X. Effects of point defects on the phase diagram of vortex states in high- T_c superconductors in $Bllc$ axis. Preprint cond-mat/0011349 at (http://xxx.lanl.gov) (2000).
- Olsson, P. & Teitel, S. Disorder driven melting of the vortex line lattice. Preprint cond-mat/0012184 at (http://xxx.lanl.gov) (2000).
- Paltiel, Y. *et al.* Instabilities and disorder-driven first-order transition of the vortex lattice. *Phys. Rev. Lett.* **85**, 3712–3715 (2000).
- Reichhardt, C., van Otterlo, A. & Zimányi, G. T. Vortices freeze like window glass: the vortex molasses scenario. *Phys. Rev. Lett.* **84**, 1994–1997 (2000).
- Fuchs, D. T. *et al.* Possible new vortex matter phases in $Bi_2Sr_2CaCu_2O_8$. *Phys. Rev. Lett.* **80**, 4971–4974 (1998).

Acknowledgements

We thank V. B. Geshkenbein for valuable discussions. This work was supported by the Israel Science Foundation – Center of Excellence Program, by the Minerva Foundation, Germany, by the Mitchell Research Fund, and by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan. D.E.F. acknowledges support from a Koshland Fellowship and an RFBR grant. P.K. and M.L. acknowledge support from the Dutch Foundation FOM. E.Z. acknowledges support from the Fundacion Antorchas – WIS program and from the Ministry of Science, Israel.

Correspondence and requests for materials should be addressed to N.A. (e-mail: nurit.avraham@weizmann.ac.il).

Coherent transfer of Cooper pairs by a movable grain

L. Y. Gorelik^{*†‡}, A. Isacson^{*†‡}, Y. M. Galperin^{†‡}, R. I. Shekhter^{*†‡} & M. Jonson^{*†‡}

^{*} Department of Applied Physics, Chalmers University of Technology; and Göteborg University, SE-412 96 Göteborg, Sweden
[†] Department of Physics, University of Oslo, PO Box 1048, N-0316 Oslo, Norway; and Division of Solid State Physics, Ioffe Institute of the Russian Academy of Sciences, St Petersburg 194021, Russia
[‡] Centre for Advanced Study, Drammensveien 78, 0271 Oslo, Norway

Superconducting circuits that incorporate Josephson junctions are of considerable experimental and theoretical interest, particularly in the context of quantum computing^{1–5}. A nanometre-sized superconducting grain (commonly referred to as a Cooper-pair box²) connected to a reservoir by a Josephson junction is an important example of such a system. Although the grain contains a large number of electrons, it has been experimentally demonstrated⁶ that its states are given by a superposition of only two charge states (differing by $2e$, where e is the electronic charge). Coupling between charge transfer and mechanical motion in nanometre-sized structures has also received considerable attention^{7–10}. Here we demonstrate theoretically that a movable Cooper-pair box oscillating periodically between two remote superconducting electrodes can serve as a mediator of Josephson coupling, leading to coherent transfer of Cooper pairs between the electrodes. Both the magnitude and the direction of the resulting Josephson current can be controlled by externally applied electrostatic fields.

First let us consider an isolated superconducting grain that is brought into direct contact with a bulk superconductor and then removed. Suppose that initially the grain was in a ground state $|n\rangle$ with a fixed number of Cooper pairs n , counted from the neutral state, while the macroscopic lead, labelled b , was in a ground state

$|\Phi_b\rangle$ with a fixed phase Φ_b of the superconducting condensate. If the switching time δt of the Josephson coupling is long enough to avoid excitations of quasiparticle states ($\delta t \Delta \ll \hbar$, where Δ is the superconducting gap and $\hbar$ is Planck's constant divided by 2π) the interaction between the grain and the lead is reduced to coherent Cooper-pair exchange (Josephson coupling). In this situation the initial state $|n\rangle|\Phi_b\rangle$ evolves into the final state $|\psi_n\rangle|\Phi_b\rangle$ where the state of the grain $|\psi_n\rangle$ can be written (as shown in the Methods):

$$|\psi_n\rangle = \sum_m s_{nm}(\Phi_b)|m\rangle = \sum_m s_{nm}(0)e^{i(n-m)\Phi_b}|m\rangle \quad (1)$$

Below we will refer to such a state as a 'Josephson hybrid'.

An important issue is that a Josephson hybrid can be created in this way only if in the contact position $\mathbf{r} = \mathbf{r}_0$ the differences between the Coulomb energies $E_n = (2en + q(\mathbf{r}))^2/2C(\mathbf{r})$ for a grain with different numbers of Cooper pairs do not exceed the Josephson energy E_J ($\mathbf{r}$ is the grain coordinate, $q(\mathbf{r})$ is the induced charge on the grain and $C(\mathbf{r})$ is the mutual capacitance). Otherwise the Coulomb blockade effect will suppress Cooper pair tunnelling and the grain will remain in the initial state. By applying gate electrodes which control the induced charge on the grain this can be achieved even if E_J tends to zero. If the gate voltage is properly chosen (say, $q(\mathbf{r}) = -e$), the Coulomb energies of the states with two subsequent numbers of Cooper pairs $n = 0, 1$ become equal during the contact time. In other words, the Coulomb blockade of tunnelling is lifted. As a result, the Josephson hybrid consists of only two states, $|0\rangle$ and $|1\rangle$. This is the single-Cooper-pair box that we shall consider. In this case, the final states of the grain $|\psi_{n=0,1}\rangle$ will be given by superpositions $|\psi_0\rangle = s_{00}|0\rangle + s_{10}e^{-i\Phi_b}|1\rangle$ and $|\psi_1\rangle = s_{11}|1\rangle + s_{01}e^{i\Phi_b}|0\rangle$. Because the Cooper-pair exchange between grain and lead decays strongly with distance, we can find the transition amplitudes s_{nm} by considering a situation when the Josephson coupling is instantaneously switched on at time $t = 0$, kept constant during a time interval t_0 (the effective contact time) and then switched off. The state of the grain, initially $|0\rangle$ or $|1\rangle$ at $t < 0$, will then evolve coherently between $|0\rangle$ and $|1\rangle$, giving rise to Rabi oscillations of the two populations with frequency $\Omega_R = \hbar^{-1}E_J$ at $t > 0$. Such oscillations were experimentally observed in ref. 6. Therefore the Josephson-hybrid amplitudes for the isolated grain will depend on the phase of the Rabi oscillations at the end of the contact time interval: $|s_{01}| = |s_{10}| = |\sin\vartheta|$ and $|s_{00}| = |s_{11}| = |\cos\vartheta|$, where $\vartheta = \Omega_R t_0$. Furthermore, a grain initially

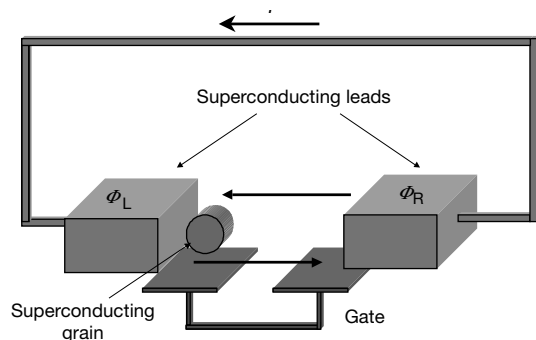


Figure 1 Schematic representation of the nanoelectromechanical system discussed in the text. A superconducting grain executes periodic motion between two superconducting bulk leads connected by a low-inductance superconducting wire. The latter facilitates a fixed phase difference between the leads. The presence of the gate ensures that the Coulomb blockade is lifted during the contacts between the grain and superconductor, which allows the grain to be in a Josephson hybrid state. As the grain moves between the leads Cooper pairs are shuttled between them creating a d.c. current through the structure.

in a hybrid state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ will, according to the superposition principle, evolve into a state $|\psi'\rangle = \alpha|\psi_0\rangle + \beta|\psi_1\rangle$ during such an event.

We now consider a grain that executes periodic motion between two superconducting leads labelled left (L) and right (R), touching them sequentially (Fig. 1). In view of the above discussion it is convenient to describe the state of the grain as $|\psi(t)\rangle = \sum_n c_n(t)e^{i\varphi_n(t)}|n\rangle$. As Josephson tunnelling occurs at each contact the single-Cooper-pair box that is executing periodic motion will mediate an indirect coherent Cooper-pair exchange between the remote leads. The total charge transmitted during a finite interval of time is determined by the whole history of the evolution of the Josephson hybrid and fluctuates strongly in general. Because the Josephson tunnelling is suppressed between the leads, the coefficients c_n remain constant while the phases φ_n evolve according to $\dot{\varphi}_n = \hbar^{-1}E_n$. In general the time evolution of the relative phases $\chi_n = \varphi_{n+1} - \varphi_n$ will depend on n . Leggett and Sols, who first addressed the question of how long the 'memory' of proximity between two remote superconductors can persist, pointed this out as an inevitable intrinsic source of destruction of such a memory¹¹. As only two charge states, characterized by a single relative phase $\chi = \varphi_1 - \varphi_0$, are involved in our Josephson hybrid, there is no room for such dephasing. However, any relaxation that is weak compared to other relevant energy scales (for example, in a real situation relaxation is caused by incoherent charge transfer owing to quasiparticle tunnelling between the grain and the leads) will remove any 'memory' of the initial conditions and the system will evolve into a stationary regime in which the grain state reproduces itself after each period. In this periodic regime we expect directed charge transfer (Josephson current) between the bulk superconductors to appear.

To find such a stationary regime we must consider the evolution of the Josephson hybrid during one period. This evolution can be presented as a sequence of independent events as shown in Fig. 2: (1) Accumulation of relative phase χ while the isolated grain moves between the electrodes. This accumulation is controlled by the electrostatic potential (gate-induced charge $q(\mathbf{r})$) between the electrodes and defined by the equation $\dot{\chi} = \hbar^{-1}(E_1 - E_0)$.

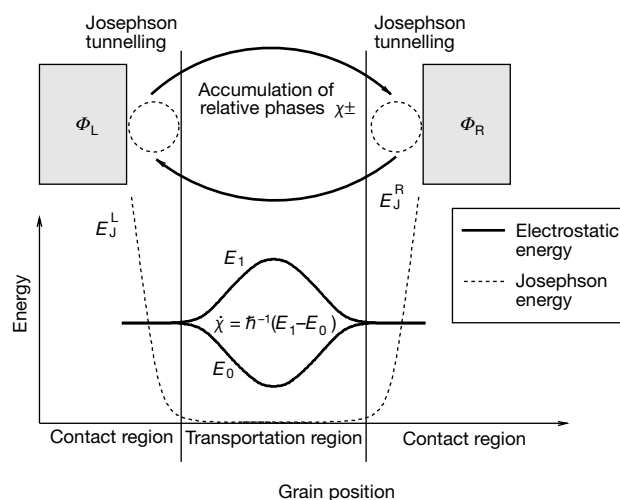


Figure 2 Illustration of the charge transport process. The grain executes periodic motion between the leads. In the contact regions the Josephson couplings $E_J^{L,R}$ are non-zero while the charging energy of the two charge states $|0\rangle$ and $|1\rangle$ are degenerate. In these regions Josephson tunnelling of Cooper pairs take place. As the grain moves into the transportation region the Josephson coupling is suppressed while the degeneracy is lifted. The relative phase χ of the Josephson hybrid now evolves according to $\dot{\chi} = \hbar^{-1}(E_1 - E_0)$. We note that the energy scales of the Josephson energy and the electrostatic energy are not the same.

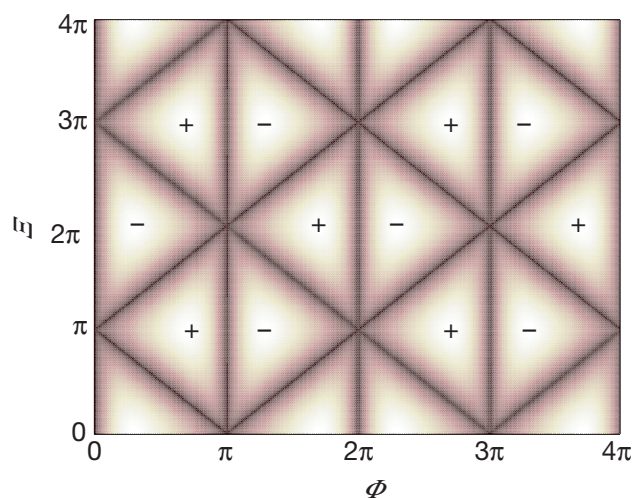


Figure 3 The magnitude of the d.c. current $\bar{I}$ in units of $I_0 = 2ef$ as a function of the phases Φ and Ξ . Regions of black correspond to no current and regions of white to $|\bar{I}/I_0| = 0.5$. The direction of the current is indicated by + or – signs. To see the triangular

structure of the current the contact time has been chosen as $\vartheta = \pi/3$. I_0 contains only the fundamental frequency of the grain's motion and the Cooper-pair charge.

(2) Transitions between states $|0\rangle$ and $|1\rangle$ as a result of Cooper-pair exchange with the bulk superconductors during a finite time interval. The phase of the probability amplitudes for these transitions is controlled by the phases $\Phi_{L,R}$ of the order parameter in the leads, while their moduli s_{nm} are determined by the contact time (see above).

It follows that in a stationary regime the evolution of the Josephson hybrid, and consequently the Josephson current, will be determined by the superconducting phases $\Phi_{L,R}$ as well as by the total relative phases χ_+ (χ_-) accumulated during the motion from the left to the right (right to left) electrode.

To calculate the Josephson current we have found a periodic solution to the quantum Liouville equation for the statistical operator $\hat{\rho}(t)$ of the system. This solution does not depend on initial conditions owing to the presence of weak relaxation caused by incoherent quasiparticle tunnelling. As shown in the Methods, the resulting d.c. current is:

$$\bar{I} = 2ef \frac{\cos\vartheta \sin^3\vartheta \sin\Phi (\cos\Xi + \cos\Phi)}{1 - (\cos^2\vartheta \cos\Xi - \sin^2\vartheta \cos\Phi)^2} \quad (2)$$

This expression contains only the fundamental frequency f of the grain's motion, the Cooper pair charge $2e$, $\Xi \equiv \chi_+ + \chi_-$ and $\Phi \equiv (\Phi_R - \Phi_L) + (\chi_+ - \chi_-)$. How the current depends on the quantities Ξ and Φ is illustrated in Fig. 3.

We observe that the d.c. current is an oscillatory function of the phase difference $\Phi_R - \Phi_L$. This is a direct manifestation of a Josephson coupling between the remote superconductors. We note that the direction of the Josephson current for a given phase difference Φ is controlled by the total relative phase Ξ accumulated during one period, which is determined by the electrostatic gate potential. Furthermore, even when the bulk superconductors have identical phases, $\Phi_R = \Phi_L$, the current does not vanish if the grain's trajectory is asymmetric, $\chi_+ \neq \chi_-$. Also, if the trajectory encloses some magnetic flux created by an external magnetic field with vector potential $\mathbf{A}(\mathbf{r})$, an additional phase $(2\pi/\Phi_0)\oint \mathbf{A}(\mathbf{r})d\mathbf{r}$ enters the expression for the phase difference Φ , which must be gauge-invariant. Here $\Phi_0 = hc/(2e)$, where c is the speed of light in vacuum, is the magnetic flux quantum.

In the above consideration we assume that the phase difference $\Phi_R - \Phi_L$ as well as the phases $\chi_{\pm}$ remain constant during the observation time. Experimentally, this means that the system should be screened from time-dependent magnetic and electric

fields. The gate electrode, as well as the insulating region between the gate and the grain, should contain as few charged dynamic defects as possible. Time-dependent switching between the states of those defects would produce charge fluctuations coupled to the charge states of the grain. Some residual dephasing should originate from charge fluctuation in the leads during contacts with the grain. The role of low-frequency fluctuations of external fields in inducing slow (compared to the oscillation period) variations of the phases $\Phi_R - \Phi_L$ and $\chi_{\pm}$ can be analysed using equation (2) by assuming the phases to be time-dependent. The current is then obtained by averaging equation (2) over the proper phase distribution. In particular, slow variations of the phase Ξ do not destroy the Cooper-pair pumping, whereas fluctuations of $\Phi_R - \Phi_L$ will suppress the current in the same fashion as in an ordinary Josephson contact.

It seems very difficult to estimate the total dephasing time theoretically. Instead, one can use the experimental results obtained by ref. 6 who observed a coherent response of a superconducting qubit during a time $\tau_{\varphi} = 10^{-9} - 10^{-8}$ s. This is a lower bound for the decoherence time in our case, because the grain is only coupled to the leads during a small part of the period of its motion. Consequently, the contact time is of the order of $\tau_c \approx T/\sqrt{l/d}$ where T is the period of the motion, d is the distance between the leads and l is the characteristic length for Josephson tunnelling, that is, $E_J^{L,R} \approx E_0 \exp(-|\delta\mathbf{r}_{L,R}|/l)$ where $|\delta\mathbf{r}_{L,R}|$ are the distances to the respective leads. It is the quantity τ_c that should be compared to the dephasing time, τ_{φ} , because during the rest of the period the grain is decoupled from the leads. For the upper bound on the dephasing time, $\tau_{\varphi} \approx 10^{-3}$ s was obtained¹² from theoretical estimates for systems fully electrically decoupled from the environment. Although the device should be carefully screened from external fields, having a movable object in the system implies that electrostatic coupling to the gate electrode as well as coupling to external degrees of freedom (phonons, photons) due to irradiation cannot be excluded. If the temperature and the quantity $\hbar\tau_c^{-1}$ are both much smaller than Δ , such coupling does not create quasiparticles and consequently does not contribute to the entropy production in the electronic subsystem. Therefore, decoherence due to particle motion can be taken into account by considering an additional force f affecting the macroscopic motion. This force may produce decoherence if it depends on the charge state $|n\rangle$ of the grain. In this case a relative shift $|\delta x| > l$ in the grain position will cause

discrimination between the different components of the hybrid during contacts with the leads and hence destroy coherence. We have estimated the dephasing time from this mechanism to be of the order of 10^{-8} s. To avoid dephasing from mechanical motion one can imagine another implementation of the Cooper pair shuttling where the alternating couplings between the grain and the leads are controlled by time-dependent potential barriers rather than by mechanical motion. In this case the relevant estimate of τ_ϕ is obtained from the experiment in ref. 6. $\square$

Methods

Derivation of equation (1)

In the case of Josephson coupling the set of conjugate variables that completely describe the system dynamics³ are $(\hat{\Phi}, \hat{N})$ and $(\hat{\phi}, \hat{n})$. Here $\hat{\Phi}(\hat{\phi})$ and $\hat{N}(\hat{n})$ are the phase and number operators of the lead (grain) condensate. In these variables the interaction between the grain and the lead is $\hat{H}_{\text{int}} = -E_J \cos(\hat{\Phi} - \hat{\phi})$ where E_J is the Josephson energy. As the initial state of the lead $|\Phi_0\rangle$ is an eigenstate of $\hat{H}_{\text{int}}$, we find $|n\rangle|\Phi_0\rangle \rightarrow \sum_m S_{mn}(\Phi_0)|m\rangle|\Phi_0\rangle$. From the commutation relation $[\hat{\Phi}, \hat{N}] = i$ it follows that $e^{-i\Phi_0 \hat{N}}|\Phi = 0\rangle = |\Phi_0\rangle$. Because the interaction conserves the total number of particles, $[\hat{n} + \hat{N}, \hat{H}_{\text{int}}] = 0$, we have $e^{-i\Phi_0 \hat{n}}|n\rangle|\Phi_0\rangle = e^{-i\Phi_0(\hat{n} + \hat{N})}|n\rangle|\Phi = 0\rangle \rightarrow \sum_m S_{mn}(0)e^{-im\Phi_0}|m\rangle|\Phi_0\rangle$.

Derivation of equation (2)

Consider a point $\mathbf{r}_A$ that the grain passes once each cycle, that is, $\mathbf{r}(t_A) = \mathbf{r}(t_A + mT) = \mathbf{r}_A$, $m = 1, 2, 3, \dots$, where T is the period of the motion. The charging term $\hat{H}_C$ and the Josephson term $\hat{H}_J$ in our hamiltonian

$$\hat{H} = \hat{H}_C + \hat{H}_J = (2e\hat{n} + q(\mathbf{r}))^2/2C(\mathbf{r}) - \sum_{s=L,R} E_J^s(\mathbf{r})\cos(\hat{\Phi}_s - \hat{\phi})$$

are never 'active' at the same time (compare Fig. 2); thus the time evolution operator $\hat{U}(t_2; t_1)$ can easily be found along with the states $|\alpha\rangle$ obeying $\hat{U}(t_A + T; t_A)|\alpha\rangle = e^{-i\epsilon_\alpha}|\alpha\rangle$. Owing to dissipation from quasiparticle tunnelling, memory of initial conditions will be lost and the system will settle into a periodic regime that is stable with respect to weak relaxation. In this case $\hat{\rho}(t)$, diagonal in the representation of quasi energy states $\hat{U}(t; t_A)|\alpha\rangle$ and periodic in time with period T , can be found by solving the quantum Liouville equation:

$$d\hat{\rho}/dt = -(i/\hbar)[\hat{H}(t), \hat{\rho}] - \nu(t)[\hat{\rho} - \hat{\rho}_0(t)] \text{ for } \nu(t) = \nu_0 \exp(-|\delta r_{L,R}(t)|/l),$$

where the last term allows for relaxation to the adiabatic equilibrium state $\hat{\rho}_0(t)$. The d.c. current is determined as the average charge delivered to the right lead:

$$\bar{I} = 2e\langle j\theta(x(t)) \rangle; \quad j = i\hbar^{-1}[\hat{H}, \hat{n}]; \quad \langle \dot{O} \rangle = \frac{1}{T} \int_T \text{Tr}[\hat{\rho}(t)\dot{O}]dt.$$

The step function $\theta(x(t))$ ensures that only tunnelling to the right lead is taken into account.

Received 1 December 2000; accepted 3 April 2001.

1. Falci, G., Fazio, R., Palma, G. M., Siewert, J. & Vedral, V. Detection of geometric phases in superconducting nanocircuits. *Nature* **407**, 355–358 (2000).
2. Averin, D. V. Adiabatic quantum computation with Cooper pairs. *Solid State Commun.* **105**, 659–664 (1998).
3. Makhlin, Yu., Schön, G. & Shnirman, A. Josephson-junction qubits with controlled couplings. *Nature* **398**, 305–307 (1999).
4. Ioffe, L. B., Geshkenbein, V. B., Feigel'man, M. V., Fauchere, A. L. & Blatter, G. Environmentally decoupled sds-wave Josephson junctions for quantum computing. *Nature* **398**, 679–781 (1999).
5. Freidman, R. J., Patel, V., Tolpygo, S. K. & Lukens, J. E. Quantum superposition of distinct macroscopic states. *Nature* **406**, 43–45 (2000).
6. Nakamura, Y., Pashkin, Yu. A. & Tsai, J. S. Coherent control of macroscopic quantum states in a single Cooper pair box. *Nature* **398**, 786–788 (1999).
7. Gorelik, L. Y. *et al.* Shuttle mechanism for charge transfer in Coulomb blockade nanostructures. *Phys. Rev. Lett.* **80**, 4526–4529 (1998).
8. Park, H. *et al.* Nanomechanical oscillations in a single C_{60} -transistor. *Nature* **407**, 57–60 (2000).
9. Tuominen, M. T., Krotkov, R. V. & Breuer, M. L. Stepwise and hysteretic transport behavior of an electromechanical charge shuttle. *Phys. Rev. Lett.* **83**, 3025–3028 (1999).
10. Erbe, A., Weiss, C., Zwirger, W. & Blick, R. H. A nanomechanical resonator shuttling single electrons at radio frequencies. Preprint cond-mat/001429 at (<http://xxx.lanl.gov>) (2000).
11. Leggett, A. J. & Sols, F. On the concept of spontaneously broken gauge symmetry in condensed matter physics. *Found. Phys.* **21**, 353–363 (1991).
12. Mooij, J. E. *et al.* Josephson persistent-current qubit. *Science* **285**, 1036–1039 (1999).

Acknowledgements

We acknowledge financial support from the Swedish Foundation for Strategic Research (SSF) through the programmes QDNS (A.I.) and ATOMICS (L.Y.G.) and from the Swedish Natural Science Research Council (NFR) (R.I.S.).

Correspondence and requests should be addressed to A.I. (e-mail: isac@fy.chalmers.se).

The complex nature of superconductivity in MgB_2 as revealed by the reduced total isotope effect

D. G. Hinks*, H. Claus*† & J. D. Jorgensen*

* Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60438, USA

† Department of Physics, University of Illinois at Chicago, Chicago, Illinois 60607, USA

Magnesium diboride, MgB_2 , was recently observed to become superconducting¹ at 39 K, which is the highest known transition temperature for a non-copper-oxide bulk material. Isotope-effect measurements, in which atoms are substituted by isotopes of different mass to systematically change the phonon frequencies, are one of the fundamental tests of the nature of the superconducting mechanism in a material. In a conventional Barden–Cooper–Schrieffer (BCS) superconductor, where the mechanism is mediated by electron–phonon coupling, the total isotope-effect coefficient (in this case, the sum of both the Mg and B coefficients) should be about 0.5. The boron isotope effect was previously shown to be large² and that was sufficient to establish that MgB_2 is a conventional superconductor, but the Mg effect has not hitherto been measured. Here we report the determination of the Mg isotope effect, which is small but measurable. The total reduced isotope-effect coefficient is 0.32, which is much lower than the value expected for a typical BCS superconductor. The low value could be due to complex materials properties, and would seem to require both a large electron–phonon coupling constant and a value of μ^* (the repulsive electron–electron interaction) larger than found for most simple metals.

The search for superconductivity in transition-metal diboride compounds is not new. In 1970, Cooper *et al.*³ were able to obtain transition temperatures (T_c) above 11 K in $\text{Zr}_{0.13}\text{Mo}_{0.87}\text{B}_2$. They also observed superconductivity in NbB_2 (ref. 3). However, as this and other searches⁴ showed, superconductivity in this class of compounds is not common, and finding superconductivity near 39 K in MgB_2 was certainly unexpected.

The presence of an isotope effect is a strong indicator of phonon mediation of the superconducting coupling. The large isotope effect measured by Bud'ko *et al.*² for B ($\alpha_B = 0.26(3)$ where (3) indicates the error in the last decimal place) clearly shows that phonons associated with B vibrations play a significant role in the superconductivity of MgB_2 . The isotope effect coefficient, α , is defined as $\alpha = -\ln T_c / \ln M$ for a one-component system, where M is the atomic mass, and as $\alpha_i = -\ln T_c / \ln M_i$ for each component, i , in a multicomponent system. The MgB_2 system is unique in that both elements have isotopes with isotopic mass differences that could

Table 1 The measured T_c s and calculated α_B s for the isotopic MgB_2 samples

B mass	Mg mass	Onset T_c (K)	10% T_c (K)	α_B onset*	α_B 10%*
10.0051	25.001	40.21(2)	40.16(2)	0.31(1)	0.31(1)
10.9952	25.001	39.06(2)	39.00(2)		
10.0051	24.305	40.23(2)	40.21(2)	0.29(1)	0.30(1)
10.9952	24.305	39.16(2)	39.09(2)		
10.0051	24.001	40.25(2)	40.20(2)	0.30(1)	0.31(1)
10.9952	24.001	39.12(2)	39.06(2)		
Average				0.30(1)	0.31(1)

* α_B is calculated for each of the Mg isotopes for the two T_c criteria.

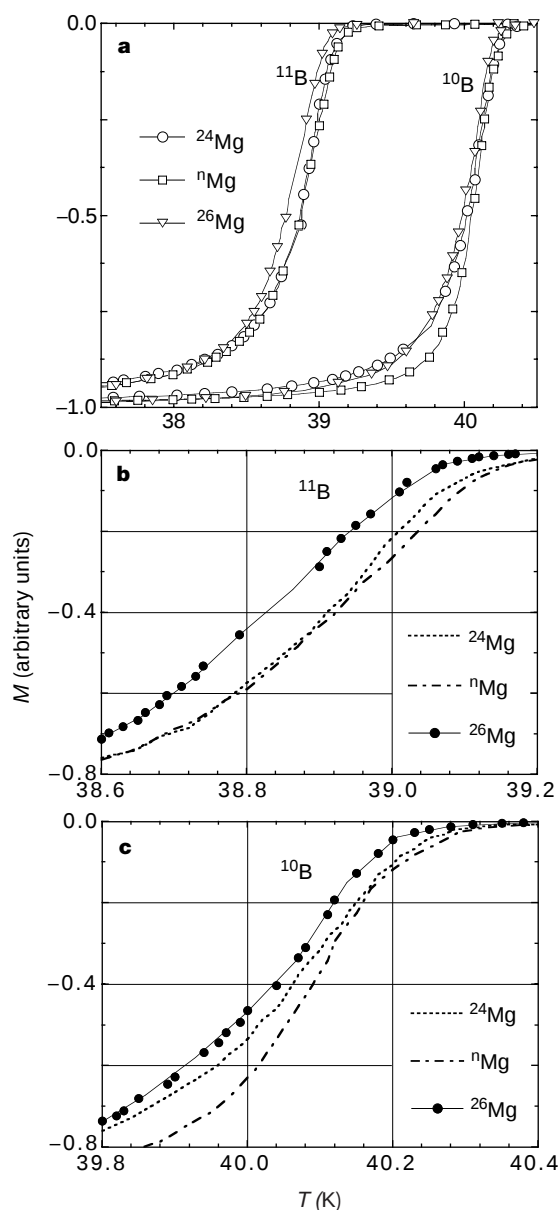


Figure 1 The superconducting transitions for the isotopically substituted MgB_2 samples. **a**, The relative magnetization for all samples. ^nMg indicates samples with natural Mg. **b,c**, The small Mg isotope effect on an expanded temperature scale for the ^{11}B and ^{10}B samples, respectively.

lead to a similar change in T_c on substitution. MgB_2 is one of the few compounds for which the total isotope coefficient, α_t , can be accurately measured.

The crystal structure of MgB_2 , composed of graphite-structured hexagonal nets of B atoms separated by hexagonal close-packed layers of Mg atoms, is electronically simple, containing only bands formed from s and p orbitals. Electronic band structure calculations by Kortus *et al.*⁵ for MgB_2 show that the Mg donates substantially both of its $3s$ electrons to the B layer, and that the Fermi surface is derived mainly from B orbitals. This leads to an ionic model of the material, $\text{Mg}^{2+}\text{B}_2^-$, and the superconductivity is, in the view of Kortus *et al.*, due to the 'metallic' boron. An and Pickett⁶ agree with the ionic model in that Mg only acts as an electron donor, but they attribute the high T_c to the two-dimensional nature of the boron Fermi surface. In their model, charge is transferred from the σ (sp^2 orbitals) bands to the π (p_z orbitals) bands to create hole

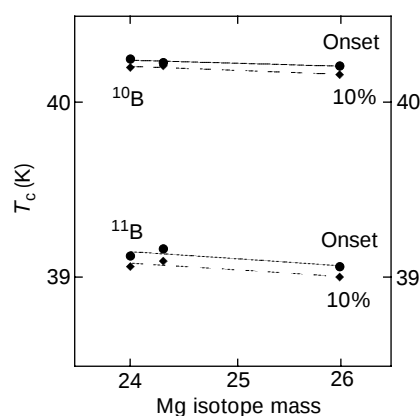


Figure 2 The superconducting transition temperatures for samples with different Mg isotopic masses for the ^{10}B and ^{11}B substituted samples. The onset T_c (solid lines) and 10% T_c (dashed lines) are least-squares fits to the data.

carriers in the strongly bonding σ bands. They conclude that the strong coupling of the holes to the in-plane B phonon lattice vibrations (E_{2g} modes) is responsible for the high T_c .

For our study, the MgB_2 samples were made from isotopic B (^{11}B and ^{10}B enriched to 99.52% and 99.49%, respectively; Eagle-Picher) and Mg isotopes (^{24}Mg and ^{26}Mg enriched to 99.92% and 99.44%, respectively; Oak Ridge). The sample size was small, about 0.1 g, because of the high cost of the isotopic Mg. The Mg (as small pieces) and B (<100 mesh powder) were weighed into small machined boron nitride (Advanced Ceramics Corp. grade HBC) crucibles with closely fitting lids. The samples were fired at 850 °C for 1.5 h under 50 bar of ultra-high-purity Ar. There was some discoloration of the BN crucibles but no major reaction. Six different samples were made, using combinations of the three Mg isotopes (24, 26 and natural) and ^{11}B and ^{10}B .

The magnetizations of the samples were measured in a non-commercial, low-field superconducting quantum interference device (SQUID) magnetometer, which was previously used extensively in the characterization of superconducting samples⁷. Approximately 5–10 mg of powder was contained in a gelatine capsule for each measurement. This magnetometer uses a Cu solenoid for fields between 0 and 100 G. With the help of a double mu-metal shield, the residual field was reduced to about 2 mG. The magnetization data were taken after cooling in zero applied field to about 20 K. A magnetic field of 1 G was then applied, and the magnetization was recorded on warming.

Figure 1a shows the magnetization data through the superconducting transitions for the six different isotopic samples. It is clear that the B isotope effect is large, and that Mg shows only a small effect. The B isotope effect was determined using both the extrapolated onset T_c and a 10% T_c criterion (that is, the temperature at which the signal is 10% of the full diamagnetic signal). Table 1 shows both the onset T_c and the 10% T_c criteria for the six samples, and the calculated values of α_B . Both sets of data agree within experimental error giving a value of 0.30(1) for α_B , within the experimental error of the value reported by Bud'ko *et al.*² of 0.26(3). The value of α_B found for MgB_2 is quite close to the B isotope effect found in the borocarbides: $\text{YNi}_2\text{B}_2\text{C}$, 0.25(4)⁸, 0.21(7)⁹, $\text{YPd}_5\text{B}_3\text{C}_{0.3}$, 0.32(4)⁸, $\text{LuNi}_2\text{B}_2\text{C}$, 0.11(5)⁹. Cheon *et al.*⁹ estimated that the total isotope effect for these materials should be near 0.5, but the lack of suitable isotopes (Y) or heavy masses (Lu) make experimental verification difficult.

The Mg isotope effect is very small, but non-zero, as illustrated in Fig. 1b and c with an expanded temperature scale. The ^{10}B and ^{11}B isotopic samples clearly show a Mg isotope effect of the expected

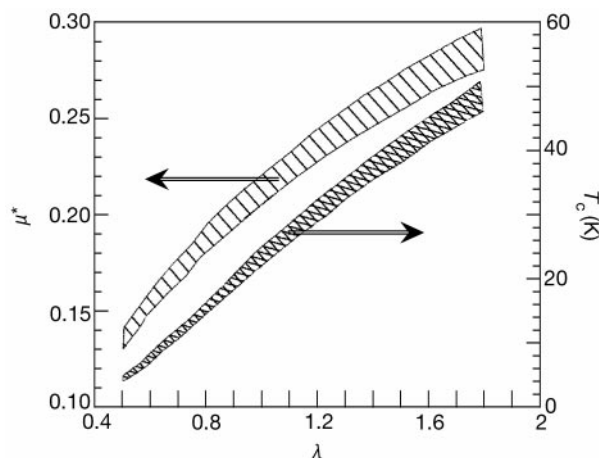


Figure 3 The allowed values of μ^* and T_c as a function of λ based on the McMillan equation if α_i ranges from 0.30 to 0.34. The T_c of MgB_2 is calculated as a function of λ using $\omega_{\text{in}} = 57.9$ meV from ref. 14.

sign; however, the natural-abundance Mg samples, labelled $^{\text{n}}\text{Mg}$, show the highest T_c s. This may be due to some small impurity content in the Mg isotopes, suppressing the T_c s of these samples relative to the natural, high-purity Mg sample. Because it is impossible to quantify the impurity effects in this measurement, all three Mg isotope masses were used for each B isotope to calculate α_{Mg} . Figure 2 shows the least-squares fit of the transition temperatures versus Mg masses for both B isotopes using both the onset and 10% T_c criteria. Averaging both the onset and 10% T_c s, and the values for both B isotopes, α_{Mg} is found to be 0.02(1).

The transition temperature for a superconductor can be written in a closed form as a product of a weighted average over the phonon frequency spectra $\langle\omega\rangle$ times a function of the coupling constant (λ) and the Coulomb repulsion pseudopotential (μ^*), and is generally approximated by the Allen-Dynes¹⁰ modification of the McMillan equation¹¹

$$T_c = \langle\omega\rangle e^{-f(\lambda, \mu^*)} \approx \frac{\omega_{\text{in}}}{1.2} e^{-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}}$$

where ω_{in} is a characteristic phonon energy determined by a weighting as defined in ref. 10. For a simple one-component, phonon-mediated BCS superconductor with $\mu^* = 0$, this form of the equation for T_c will give $T_c \propto \omega_{\text{in}} \propto (1/M)^{1/2}$ because λ is, in general, thought to be mass independent. The isotope coefficient under these assumptions is 1/2. However, it may vary greatly if $\mu^* \neq 0$ and λ is small (as is, for example, the case for some low- T_c materials). Garland¹² has shown that for transition metals, Coulomb effects reduce α due to the presence of d orbitals whereas sp metals generally show a nearly full isotope effect. Using the approximate form of the McMillan equation shown above, α can be written^{11,13} in a closed form as a function of λ and μ^* ,

$$\alpha = \frac{1}{2} \left[1 - \frac{1.04(1+\lambda)(1+0.62\lambda)\mu^{*2}}{\lambda - \mu^*(1+0.62\lambda)^2} \right]$$

which reduces to 1/2 for $\mu^* = 0$.

In the case of a multicomponent system, α_i , the total isotope effect coefficient, which is the sum of all individual isotopic effect coefficients, would tend to 1/2 at low values of μ^* and large values of λ (ref. 13). In this case, the functional form of the averaging process over the phonon density of states to obtain $\langle\omega\rangle$ will determine the individual isotope effect coefficients, α_i . Recently, Osborn *et al.*¹⁴ have calculated α_{B} and α_{Mg} using a Born-von-Karman model of the measured phonon density of states. Depending on their

weighting of the phonon density of states, they obtain 0.42 and 0.08 or 0.28 and 0.22 for α_{B} and α_{Mg} , respectively. As expected, the individual isotope-effect coefficients varied widely depending on the exact functional form of the prefactor, $\langle\omega\rangle$, for the T_c equation; however, the total isotope coefficient remained at 1/2.

MgB_2 appears to be a conventional, BCS phonon-mediated sp superconductor with an intermediate-to-strong electron-phonon coupling strength. Specific heat measurements¹⁵ and ^{11}B nuclear spin lattice relaxation measurements¹⁶ indicate a moderately strong coupling superconductor. The superconducting gaps determined from tunnelling measurements range from 2 meV to 7 meV (refs 17–20), spanning the range of weak to moderate coupling strength (5.9 meV is the weak coupling limit gap value). There is some evidence that the low gaps result from a weakened surface layer¹⁹, thus, if the largest gaps are intrinsic to the material, tunnelling measurements would place MgB_2 in the moderately strong limit. Estimates of λ range from 0.65 to 1.2 with values of μ^* ranging from near 0 to 0.2 (refs 5,14,21).

The measured reduced total isotope coefficient of 0.32 could be the result of a large Coulomb pseudopotential. Our measured value can place some limits on the values of λ and μ^* . Figure 3 shows the allowable range of μ^* for different λ s calculated with the equation shown above for α . The value of μ^* must, of necessity, be larger than the canonical value of 0.1 generally assumed for this variable in order to account for the reduced total isotope coefficient. The results shown in Fig. 3 place demanding constraints on theoretical attempts to explain the high T_c . For example, the recent electronic structure calculations for MgB_2 by Kong *et al.*²¹ resulted in a value of $\lambda \approx 0.85$, requiring a value of μ^* of about 0.1 to account for the observed T_c of 39 K. These values of λ and μ^* would lead to a total isotope coefficient near 0.5, inconsistent with the measured value. T_c can also be calculated if the value of ω_{in} is known. Using the value of 57.9 meV determined by Osborn *et al.*¹⁴, the calculated T_c s as a function of λ are also shown in Fig. 3. The values found for λ and μ^* for a T_c of 39 K, about 1.4 and 0.25 respectively, are larger than most estimates so far determined for these parameters.

Our measurement of the Mg and B isotope effects shows that there is substantial coupling of B atom vibrations to the electronic structure. The very small value of 0.02(1) for α_{Mg} shows that vibrational frequencies of Mg atoms have almost no effect on T_c ; however, we cannot rule out the possibility that the Mg vibration modes still contribute to T_c . The presence of an isotope effect indicates a phonon contribution to T_c , but an absence—or the measurement of a small value—is ambiguous. It is only the total

isotope effect that has any quantitative meaning within the context of simple BCS superconductivity.

We have argued that our observed reduced total isotope coefficient is the result of a large Coulomb repulsion resulting in a large value of μ^* . Certainly, many other causes for a reduced total isotope effect can occur—for example, anharmonic lattice vibration could affect the mass dependence of $\langle\omega\rangle$, or complex dimensional effects on the Fermi surface could affect $f(\lambda, \mu^*)$. The model of An and Pickett⁶, for example, might lead to a reduced total isotope effect if the two-dimensional nature of the hole band is found to have an effect on λ or μ^* . This difference between the measured total isotope effect and 1/2 may appear insignificant, but the cause of this discrepancy may be in some way related to the high T_c observed in this material. Although it is possible to calculate T_c for MgB₂ based on reasonable choices of parameters using the McMillan equation, correctly modelling α_t (and α_{Mg} and α_B) may place constraints on these calculations that will provide important insight into how such a high T_c is achieved in this simple *sp* metal.

Note added in proof: The reduced isotope effect in MgB₂ has recently been explained²² as being due to the large anharmonicity of the E_{2g} mode. If correct, this would reduce the value of λ and μ^* required to account for the high T_c of the material. □

Received 8 March; accepted 25 April 2000.

1. Nagamatsu, J., Nakagawa, N., Muraoka, T., Zenitani, Y. & Akimitsu, J. Superconductivity at 39 K in magnesium diboride. *Nature* **410**, 63–64 (2001).
2. Bud'ko, S. L. *et al.* Boron isotope effect in superconducting MgB₂. *Phys. Rev. Lett.* **86**, 1877–1880 (2001).
3. Cooper, A. S., Korenzwit, E., Longinotti, L. D., Matthias, B. T. & Zachariasen, W. H. Superconductivity: the transition temperature peak below four electrons per atom. *Proc. Natl Acad. Sci. USA* **67**, 313–319 (1970).
4. Leyarovska, L. & Leyarovski, E. A search for superconductivity below 1 K in transition metal borides. *J. Less Common Met.* **67**, 249–255 (1979).
5. Kortus, J., Mazin, I. L., Belashchenko, K. D., Antropov, V. P. & Boyer, L. L. Superconductivity of metallic boron in MgB₂. Preprint cond-mat/0101446 at (<http://xxx.lanl.gov>) (2001).
6. An, J. M. & Pickett, W. E. Superconductivity of MgB₂ from hole-doped covalent bonds. Preprint cond-mat/0102391 at (<http://xxx.lanl.gov>) (2001).
7. Kostic, P. *et al.* Paramagnetic Meissner effect in Nb. *Phys. Rev. B* **53**, 791–801 (1996).
8. Lawrie, D. D. & Franck, J. P. Boron isotope effect in Ni and Pd based borocarbide superconductors. *Physica C* **245**, 159–163 (1995).
9. Cheon, K. O., Fisher, I. R. & Canfield, P. C. Boron isotope effect in single-crystal YNi₂B₂C and LuNi₂B₂C superconductors. *Physica C* **312**, 35–39 (1999).
10. Allen, P. B. & Dynes, R. C. Transition temperature of strong-coupled superconductors reanalyzed. *Phys. Rev. B* **12**, 905–922 (1975).
11. McMillan, W. L. Transition temperature of strong-coupled superconductors. *Phys. Rev.* **167**, 331–344 (1968).
12. Garland, J. W. Jr. Isotope effect in superconductivity. *Phys. Rev. Lett.* **11**, 114–119 (1963).
13. Carbotte, J. P. Properties of boson-exchange superconductors. *Rev. Mod. Phys.* **62**, 1027–1157 (1990).
14. Osborn, R., Goremychkin, E. A., Kolesnikov, A. I. & Hinks, D. G. Phonon density-of-states in MgB₂. *Phys. Rev. Lett.* (in the press); also as preprint cond-mat/0103064 at (<http://xxx.lanl.gov>) (2001).
15. Kremer, R. K., Gibson, B. J. & Ahn, K. Heat capacity of MgB₂: Evidence for moderately strong coupling behavior. Preprint cond-mat/0102432 at (<http://xxx.lanl.gov>) (2001).
16. Koteagawa, H., Ishida, K., Kitaoka, Y., Muranaka, T. & Akimitsu, J. Evidence for strong-coupling s-wave superconductivity in MgB₂: ¹B NMR study. Preprint cond-mat/0102334 at (<http://xxx.lanl.gov>) (2001).
17. Rubio-Bollinger, G., Suederow, H. & Vieira, S. Tunneling spectroscopy in small grains of superconducting MgB₂. Preprint cond-mat/0102242 at (<http://xxx.lanl.gov>) (2001).
18. Karapetrov, G., Iavarone, M., Kwok, W. K., Crabtree, G. W. & Hinks, D. G. Scanning tunneling spectroscopy in MgB₂. *Phys. Rev. Lett.* (in the press); also as preprint cond-mat/0102312 at (<http://xxx.lanl.gov>) (2001).
19. Schmidt, H., Zasadzinski, J. F., Gray, K. E. & Hinks, D. G. Energy gap from tunneling and metallic Shrivin contacts into MgB₂: evidence for a weakened surface layer. *Phys. Rev. B* (in the press); also as preprint cond-mat/0102389 at (<http://xxx.lanl.gov>) (2001).
20. Sharoni, A., Felner, I. & Millo, O. Tunneling spectroscopy and magnetization measurements of the superconductor properties of MgB₂. *Phys. Rev. B* (in the press); also as preprint cond-mat/0102325 at (<http://xxx.lanl.gov>) (2001).
21. Kong, Y., Dolgov, O. V., Jepsen, O. & Anderson, O. K. Electron-phonon interaction in the normal and superconducting states of MgB₂. *Phys. Rev. B* (in the press); also as preprint cond-mat/0102499 at (<http://xxx.lanl.gov>) (2001).
22. Yildirim, T. *et al.* Giant anharmonicity and non-linear electron-phonon coupling in Mg B₂: combined first principles calculations and neutron scattering study. Preprint cond-mat/0103469 at (<http://xxx.lanl.gov>) (2001).

Acknowledgements

This work was supported by the US Department of Energy, Office of Science.

Correspondence and requests for materials should be addressed to D.G.H. (e-mail: hinks@anl.gov).

Constraints on hydrothermal processes and water exchange in Lake Vostok from helium isotopes

Philippe Jean-Baptiste*, Jean-Robert Petit†, Vladimir Ya. Lipenkov‡, Dominique Raynaud† & Nartsiss I. Barkov‡

* Laboratoire des Sciences du Climat et de l'Environnement, CEA/CNRS, Centre d'études de Saclay, 91191 Gif sur Yvette, France

† Laboratoire de Glaciologie et de Géophysique de l'Environnement, CNRS, Grenoble, 38402 Saint Martin d'Hères, France

‡ Arctic and Antarctic Research Institute, Beringa Street 38, 199397 St Petersburg, Russia

Lake Vostok, the largest subglacial lake in Antarctica, is covered by the East Antarctic ice sheet, which varies in thickness between 3,750 and 4,100 m (ref. 1). At a depth of 3,539 m in the drill hole at Vostok station, sharp changes in stable isotopes and the gas content of the ice delineate the boundary between glacier ice and ice accreted through re-freezing of lake water². Unlike most gases, helium can be incorporated into the crystal structure of ice during freezing³, making helium isotopes in the accreted ice a valuable source of information on lake environment. Here we present helium isotope measurements from the deep section of the Vostok ice core that encompasses the boundary between the glacier ice and accreted ice, showing that the accreted ice is enriched by a helium source with a radiogenic isotope signature typical of an old continental province. This result rules out any significant hydrothermal energy input into the lake from high-enthalpy mantle processes, which would be expected to produce a much higher ³He/⁴He ratio. Based on the average helium flux for continental areas, the helium budget of the lake leads to a renewal time of the lake of the order of 5,000 years.

In the upper part of the Antarctic ice sheet, helium is of atmospheric origin⁴. Previous measurements obtained between 1989 and 1992 at various depths—from 116 to 2,501 m (Table 1)—in the Vostok ice core indicate a nearly constant helium concentration: [He] = 12.1 ± 1.3 nmol kg⁻¹, with an isotopic ratio $R = (0.97 \pm 0.03)R_a$, where $R = {}^3\text{He}/{}^4\text{He}$ and R_a is the atmospheric ratio (1.38×10^{-6}). The measured concentration represents only a fraction of the helium initially present in the air bubbles at the pores' close-off (21.4–22.3 nmol kg⁻¹ at Vostok⁴) owing to the upward diffusion of helium to the atmosphere as the air bubbles are increasingly compressed⁵. In deeper layers, the helium isotope distribution may be influenced by the radiogenic signature from the bedrock⁶. In Lake Vostok, dissolved helium may therefore have three main possible origins: (1) atmospheric helium, with an isotope ratio close to R_a , trapped in glacier ice and released into the lake water as the basal layers are melted; (2) radiogenic crustal helium produced by the decay of U and Th in the bedrock and bottom sediments, characterized by $R < 0.05R_a$ (refs 7,8); and (3) mantle helium with $R > 8R_a$ (refs 8–10). (We note that, because of the large ice–gas age difference at pore close-off in central Antarctica, of the order of 1,000 years or greater, there should be no imprint of any tritiogenic ³He produced by the decay of natural tritium.)

We analysed 14 new samples from the Vostok ice core from a depth range of 3,530–3,608 m (Table 1), which encompasses the transition between glacier ice and the refrozen lake water at 3,539 m depth. The main feature of the helium isotope distribution with depth (Fig. 1) is the pronounced difference in both the helium concentration and isotope ratio between glacier ice and the refrozen lake water. For depths between 3,530 and 3,535 m, the mean helium concentration, [He] = 10.0 ± 1.3 nmol kg⁻¹, is similar to

that measured in the upper layers of the Vostok ice core, with $R = (1.03 \pm 0.08)R_a$, close to the atmospheric value. Below 3,539 m depth, the helium content of the samples is enriched by a factor of 3 with respect to glacier ice ($[He] = 33.9 \pm 2.6 \text{ nmol kg}^{-1}$) and exhibits a clear radiogenic signature, $R = (0.25 \pm 0.04)R_a$, typical of old stable continental provinces.

Lake Vostok's ice ceiling is tilted, with an ice thickness of 4,300 m in the north and 3,750 m in the south. This north–south gradient of the pressure-dependent melting point favours glacier ice melting in the north and lake water refreezing in the south^{11,12}, where the Vostok station is located. Radar studies also suggest basal melting on the western margin upstream from Vostok¹¹; however, the numerous solid inclusions observed in the upper part of the lake ice in the Vostok core indicate that freezing may indeed occur as soon as glacial ice loses contact with bedrock². Water influx also occurs as a result of ice melting off the base of the glacier over the lake catchment in response to geothermal flux and shear-strain heating.

The long-term mass balance between the input of melt water and the accreted ice export allows us to determine the helium isotope budget of the lake. Per unit of time, the amount of helium released to the lake by the melted glacier ice (with a helium concentration C_{He1}^{in}) and basal water runoff (with a helium concentration C_{He2}^{in}), together with the amount of crustal helium injected into the lake from the bedrock (Q_{He}), must be balanced by the amount of helium that leaves the lake by direct diffusion through the glacier ice/lake ice interface (Q_{He}^{dif}), together with the exported lake ice of helium concentration C_{He}^{out} . This leads to the following equations written for ^4He and ^3He , respectively:

$$(1 - \alpha)mC_{He1}^{\text{in}} + \alpha mC_{He2}^{\text{in}} + Q_{He} = mC_{He}^{\text{out}} + Q_{He}^{\text{dif}} \quad (1)$$

$$(1 - \alpha)mR_1^{\text{in}}C_{He1}^{\text{in}} + \alpha mR_2^{\text{in}}C_{He2}^{\text{in}} + RQ_{He} = mR^{\text{out}}C_{He}^{\text{out}} + Q_{He}^{\text{dif}}R \quad (2)$$

where m is the amount of water/ice exchanged per unit of time, α is the water runoff fraction, and R_1^{in} , R_2^{in} and R^{out} are the isotope ratios of melted glacier ice, runoff basal water, and accreted ice respectively. R is the isotope ratio of the helium source Q_{He} , and Q_{He}^{dif} and Q_{He}^{out} are the amounts of ^4He and ^3He leaving the lake by direct

diffusion through the glacier ice/lake ice interface.

Q_{He}^{dif} is negligible because the ^3He concentration gradient between the lake ice and the glacier ice is zero within the experimental uncertainties. Hence, combining equations (1) and (2), we can deduce R by:

$$R = [R_{\text{out}}C_{He}^{\text{out}} - (1 - \alpha)R_1^{\text{in}}C_{He1}^{\text{in}} - \alpha R_2^{\text{in}}C_{He2}^{\text{in}}] / [C_{He}^{\text{out}} + (Q_{He}^{\text{dif}}/m) - (1 - \alpha)C_{He1}^{\text{in}} - \alpha C_{He2}^{\text{in}}] \quad (3)$$

C_{out} and R_{out} can be taken as the mean values in the accreted ice below the interface ($C_{He}^{\text{out}} = 33.9 \pm 2.6 \text{ nmol kg}^{-1}$ and $R_{\text{out}} = (0.25 \pm 0.04)R_a$). The glacier ice layer immediately above the interface, between 3,530 and 3,535 m depth, is assumed to be representative of the input of melted glacier ice occurring in the northern part of the lake ($C_{He1}^{\text{in}} = 10.0 \pm 1.3 \text{ nmol kg}^{-1}$ and $R_1^{\text{in}} = (1.03 \pm 0.08)R_a$). Although this layer shows no crustal helium tagging, its near-bottom origin is suggested by the occurrence of small size moraine particles at the base of the section: our preliminary study of the dust in the 3,400–3,538 m layer does indeed reveal the presence of particles, 4–20 μm in diameter, which exceed in size the aeolian dust found at Vostok and are interpreted as crustal dust created by glacier work on the western side of the lake. Crustal helium tagging of the basal layer depends on the intensity of the helium flux and on the residence times of the ice resting on the bedrock. We note that the lack of helium crustal signature in a glacier basal layer was also observed in Greenland. The lowermost 200 m of the GISP2 core⁶ shows two sharp spikes of crustal helium, but these spikes do not correspond to layers of silty ice, although there is such a layer at the bottom of the core. The author of ref. 6 concludes that the lack of a simple exponential helium profile at the bottom of the core shows that the crustal helium injection does not involve ice resting directly on the bedrock but may represent disturbed ice layers which have had a significantly longer residence time above the bedrock—long enough to accumulate a large amount of crustal helium from below. The helium isotope values of these crustal anomalies⁶ ($C_{He2}^{\text{in}} \approx 20 \text{ nmol kg}^{-1}$ and $R_2^{\text{in}} \approx 0.02R_a$) are taken as representative of the runoff water that melts at the ice–bedrock interface.

With the above data set and neglecting Q_{He}^{dif}/m , equation (3) leads to a maximum isotope ratio R of helium escaping from the lake bottom of $R = (0.11 \pm 0.05)R_a$ for values of α of up to 50%. Based on a basal melt rate of a few millimetres per year for the water runoff^{1,13}, a hydrological catchment area equal to several times the lake's surface area, and a melt rate of 100 mm yr^{-1} in the northern half of the lake¹¹, the most probable value for α falls in the range 5–15%, thus leading to an isotope ratio $R < 0.06R_a$. This low $^3\text{He}/^4\text{He}$

Table 1 Helium isotope data

Depth (m)	$[^4\text{He}]$ (nmol kg^{-1})	R/R_a
Previous sampling (1989–92)*		
116–165	10.4 ± 1.2	0.985 ± 0.011
120–145	11.9 ± 1.4	1.017 ± 0.031
1,252–1264	13.3 ± 0.5	0.943 ± 0.007
2,441–2501	13.2 ± 0.9	0.954 ± 0.015
Deep core sampling (1997)		
3,530	9.5 ± 0.8	1.111 ± 0.005
3,531	8.7 ± 0.8	1.098 ± 0.005
3,535†	9.8 ± 0.8	0.951 ± 0.005
3,535†	11.9 ± 0.8	0.945 ± 0.005
3,540	29.1 ± 0.8	0.282 ± 0.005
3,551	38.4 ± 0.8	0.262 ± 0.005
3,565	36.0 ± 0.8	0.225 ± 0.005
3,568	35.8 ± 0.8	0.235 ± 0.005
3,576	32.1 ± 0.8	0.253 ± 0.005
3,578	34.5 ± 0.8	0.240 ± 0.005
3,580	35.8 ± 0.8	0.327 ± 0.005
3,599	33.1 ± 0.8	0.191 ± 0.005
3,605	31.8 ± 0.8	0.267 ± 0.005
3,608	32.5 ± 0.8	0.219 ± 0.005

Each sample consisted of one piece of ice, typically 10 g, cut from the core immediately following its retrieval (immediate processing of the sample is essential to avoid helium post-coring diffusion⁴). The samples were stored in one inch diameter copper cylinders evacuated for 2 min with a primary pump to 10^{-2} – 10^{-3} torr and sealed with a metal clamp (the helium blank due to air contamination is negligible, $< 4.5 \times 10^{-14}$ mol). In spite of the strict sampling protocol, some helium is lost during sample preparation. With a total preparation time of 15 min (including 2 min pumping time), we calculate that this post-coring loss could amount to 10%, but with no significant isotope fractionation. Samples were analysed at LSCE (Saclay) using standard procedures²¹.

* Unpublished data.

† Duplicate samples.

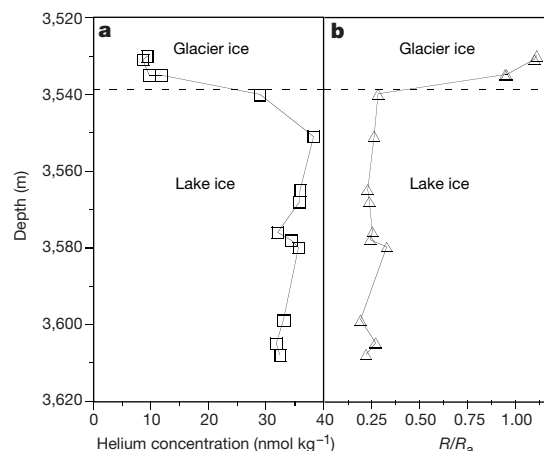


Figure 1 Measured vertical distribution of helium isotopes across the glacier ice/lake ice boundary (dotted line) at Vostok. **a**, Vertical profile of helium concentration. **b**, Helium isotope ratio $R = ^3\text{He}/^4\text{He}$, scaled to the atmospheric ratio R_a .

ratio is typical of radiogenic crustal helium^{6,7}. Taking into account the diffusive term $Q_{\text{He}}^{\text{dif}}/m$ leads to an even lower estimate. $^3\text{He}/^4\text{He}$ is a sensitive indicator of mantle involvement in crustal thermal activity¹⁴. In continental areas of recent tectonic-magmatic activity such as the Yellowstone¹⁵ or Larderello¹⁶ geothermal areas, $^3\text{He}/^4\text{He}$ is at its highest, accompanied by excess heat flow. On the contrary, in stable areas, where low $^3\text{He}/^4\text{He}$ ratios are found, the mantle contribution to the heat flow is also low and reaches a minimum value that can be considered as the continental background⁸. Therefore, our data virtually rule out any mantle-derived high-enthalpy hydrothermal manifestations at the bottom of the lake.

Rewriting equation (1) using $m = M/\tau$, where τ is the renewal time of the lake and M its mass (with $M = \rho HS$ where ρ is the lake mean density, H the mean depth and S the average surface area), we obtain:

$$\Phi_{\text{He}} = (C_{\text{He}}^{\text{out}} - (1 - \alpha)C_{\text{He}}^{\text{in}} - \alpha C_{\text{He}}^{\text{in}})\rho H/\tau + \Phi_{\text{He}}^{\text{dif}} \quad (4)$$

where Φ_{He} and $\Phi_{\text{He}}^{\text{dif}}$ are the helium flux from the bedrock and the helium diffusive flux respectively, scaled to the average lake surface area S ($\Phi_{\text{He}} = Q_{\text{He}}/S$ and $\Phi_{\text{He}}^{\text{dif}} = Q_{\text{He}}^{\text{dif}}/S$). The diffusive component $\Phi_{\text{He}}^{\text{dif}}$ can be estimated from the helium concentration gradient across the glacier ice/lake ice interface (∇C) and the helium diffusion coefficient (D_{He}) by $\Phi_{\text{He}}^{\text{dif}} = D_{\text{He}} \times \nabla C$. Taking for ∇C the concentration gradient measured at Vostok across the glacier ice/lake ice boundary ($\nabla C \approx 3.8 \mu\text{mol m}^{-4}$), and a diffusion coefficient $D_{\text{He}} = 5 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ (refs 3, 17, 18) at the temperature of the interface ($T = -3^\circ\text{C}$), we calculate a diffusive flux of $60 \text{ nmol m}^{-2} \text{ yr}^{-1}$. If we adopt for Φ_{He} the average estimate for the continental crust of $1.46 \mu\text{mol m}^{-2} \text{ yr}^{-1}$ assuming that the helium loss from crust is equal to production¹⁴, it follows that the first term of equation (4) (advective component) equals $1.4 \mu\text{mol m}^{-2} \text{ yr}^{-1}$, leading to a residence time τ ranging from 4,800 to 5,000 years (taking α between 5% and 15% and a mean lake depth of 300 m; ref. 19).

This residence time corresponds to a lake mean melting/accretion rate $h = 2(1 - \alpha)H/\tau$ of 11 cm yr^{-1} (if the respective surface area of melting and freezing are assumed to be roughly equal). This value is of the same order of magnitude as that inferred from airborne 60-MHz radar transects above the lake¹¹. Owing to the low heat flux measured at Vostok, corresponding to a maximum apparent accretion rate of 4 mm yr^{-1} (ref. 20), this high accretion rate implies that upon freezing, most of the latent heat is released within the lake water itself rather than in the overlying ice. This is consistent with the recent hypothesis of Souchez *et al.*¹², who suggest that the main mechanism responsible for lake ice formation is loose frazil ice crystals produced in supercooled waters rising along the tilted ice–water interface.

In reasonable agreement with geophysical inferences¹¹ indicating the significant water exchange between the base of the ice sheet and the lake waters, helium data in the deep Vostok ice core point to appreciable mass exchange occurring between the subglacial lake and the overlying ice sheet. More importantly, the helium isotope data virtually eliminates any significant high-enthalpy hydrothermal input to the energy budget available for metabolic activity within the lake, an important issue with regards to the sustainability of life in this isolated environment. □

Received 1 December 2000; accepted 13 March 2001.

- Kapitsa, A. P., Ridley, J. K., Robin, G. Q., Siegert, M. J. & Zotikov, I. A. A large deep freshwater lake beneath the ice of central East Antarctica. *Nature* **381**, 684–686 (1996).
- Jouzel, J. *et al.* More than 200 meters of lake ice above subglacial lake Vostok, Antarctica. *Science* **286**, 2138–2141 (1999).
- Haas, J., Bullemer, B. & Kahane, A. Diffusion de l'hélium dans la glace monocristalline. *Solid State Commun.* **9**, 2033–2035 (1971).
- Jean-Baptiste, P., Raynaud, D., Mantis, F., Sowers, T. & Barkov, N. Measurement of helium isotopes in Antarctic ice: preliminary results from Vostok. *CR Acad. Sci. Paris* **316**, 491–497 (1993).
- Craig, H. & Chou, C. C. Helium isotopes and gases in Dye 3 ice cores. *Eos* **63**, 298 (1982).
- Craig, H. in *Abstr. Int. Conf. On Stable Isotopes and Isotope Effects* (eds Soulié, E. & Roth, E.) 1–112 (Commissariat à l'Energie Atomique, Paris, 1999).
- Andrews, J. N. The isotopic composition of radiogenic helium and its use to study groundwater movement in confined aquifers. *Chem. Geol.* **49**, 339–351 (1985).

- Mamyrin, B. A. & Tolstikhin, I. N. *Helium Isotopes in Nature* (Elsevier, Amsterdam, 1984).
- Craig, H. & Lupton, J. E. Primordial neon, helium and hydrogen in oceanic basalts. *Earth Planet. Sci. Lett.* **31**, 369–385 (1976).
- Farley, K. A. & Neroda, E. Noble gases in the Earth's mantle. *Annu. Rev. Earth Planet. Sci.* **26**, 189–218 (1998).
- Siegert, M. J., Kwok, R., Mayer, C. & Hubbard, B. Water exchange between the subglacial Lake Vostok and the overlying ice sheet. *Nature* **403**, 643–646 (2000).
- Souchez, R., Petit, J. R., Tison, J. L., Jouzel, J. & Verbeke, V. Ice formation in subglacial Lake Vostok, Central Antarctica. *Earth Planet. Sci. Lett.* **181**, 529–538 (2000).
- Savvin, A., Greve, R., Calov, R., Mugge, B. & Hutter, K. Simulation of the Antarctic ice sheet with a three-dimensional polythermal ice-sheet model, in support of the EPICA project. II. Nested high resolution treatment of Dronning Maud Land, Antarctica. *Ann. Glaciol.* **30**, 69–75 (2000).
- O'Nions, R. K. & Oxburgh, E. R. Heat and helium in the earth. *Nature* **306**, 429–431 (1983).
- Craig, H., Lupton, J. E., Welhan, J. A. & Poreda, R. Helium isotope ratios in Yellowstone and Lassen Park volcanic gases. *Geophys. Res. Lett.* **5**, 897–900 (1978).
- Hooker, P. J., Bertrami, R., Lombardi, S., O'Nions, R. K. & Oxburgh, E. R. Helium-3 anomalies and crust–mantle interaction in Italy. *Geochim. Cosmochim. Acta* **49**, 2505–2513 (1985).
- Davy, J. G. & Miller, K. W. The diffusion of helium through ice. *Solid State Commun.* **3**, 1459–1461 (1970).
- Jones, S. J. Diffusion of helium in ice. *Phys. Canada* **27**, 60–61 (1971).
- Siegert, M. J. Antarctic subglacial lakes. *Earth Sci. Rev.* **50**, 29–50 (2000).
- Salamatin, A., Vostretsov, R. N., Petit, J. R., Lipenkov, V. Y. & Barkov, N. I. Geophysical and paleoclimatic implications of the temperature profile from the deep borehole at Vostok station (Antarctica). *Data Glaciol. Stud.* **85**, 233–240 (1998).
- Jean-Baptiste, P., Mantis, F., Dapigny, A. & Stievenard, M. Design and performance of a mass spectrometric facility for measuring helium isotopes in natural waters and for low-level tritium determination by the ^3He ingrowth method. *Appl. Radiat. Isot.* **43**, 881–891 (1992).

Acknowledgements

Vostok is a joint project between Russia, France and the USA. We thank the Russian Antarctic Expedition for station and drilling-operation support, and the IFRTF (Institut Français de Recherche et Technologies Polaires) and the NSF Office of Polar Program for logistic support. The project was supported in Russia by the Russian Ministry of Sciences, and in France by CNRS and CEA. We thank A. Salamatin, P. Duval and J. Meyssonier for discussions.

Correspondence should be addressed to P.J.-B. (e-mail: pjb@lsce.saclay.cea.fr).

Earthquake triggering by seismic waves following the Landers and Hector Mine earthquakes

J. Gombert*, P. A. Reasenbergt, P. Bodin† & R. A. Harris‡

* US Geological Survey, Center for Earthquake Research and Information, 3876 Central Avenue Suite 2, Memphis, Tennessee 38152-3050, USA

† US Geological Survey, MS 977, 345 Middlefield Road, Menlo Park, California 94025, USA

‡ Center for Earthquake Research and Information, The University of Memphis, 3876 Central Avenue, Suite 1, Memphis, Tennessee 38152-3050, USA

The proximity and similarity of the 1992, magnitude 7.3 Landers and 1999, magnitude 7.1 Hector Mine earthquakes in California permit testing of earthquake triggering hypotheses not previously possible. The Hector Mine earthquake confirmed inferences that transient, oscillatory 'dynamic' deformations radiated as seismic waves can trigger seismicity rate increases, as proposed for the Landers earthquake^{1–6}. Here we quantify the spatial and temporal patterns of the seismicity rate changes⁷. The seismicity rate increase was to the north for the Landers earthquake and primarily to the south for the Hector Mine earthquake. We suggest that rupture directivity results in elevated dynamic deformations north and south of the Landers and Hector Mine faults, respectively, as evident in the asymmetry of the recorded seismic velocity fields. Both dynamic and static stress changes seem important for triggering in the near field with dynamic stress changes dominating at greater distances. Peak seismic velocities recorded for each earthquake suggest the existence of, and place bounds on,

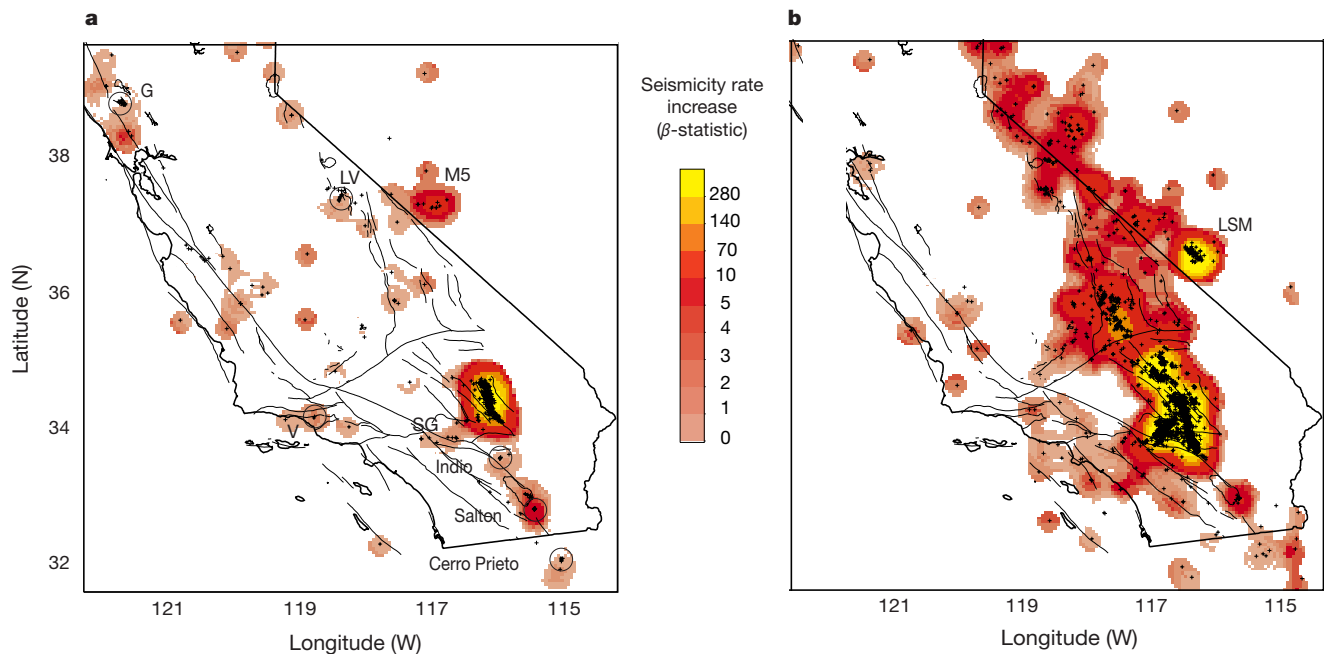


Figure 1 Seismicity rate increases associated with Hector Mine and Landers mainshocks. Average increases are quantified with the β -statistic⁷ calculated for spatially smoothed (20 km) seismicity ($M \geq 2.0$) catalogued by the Council for a National Seismic System (CNSS) using data from the Southern California Seismic Network (SCSN) and Northern California Seismic Network (NCSN). Rate change is calculated for a two-week period immediately after each mainshock, relative to the background period 1987–1992. Plus signs indicate earthquakes in the period after the mainshock. Areas with a β -statistic ≥ 2.0 are suggestive of a significant average rate increase, but the concurrence of the increase with the Hector Mine earthquake must be confirmed by examining seismicity

time histories. White areas, no catalogued earthquakes in post- and pre-mainshock periods or no detectable rate increase; thin lines, surface fault traces. **a**, Hector Mine response. Circles mark sites of triggered activity. LV, Long Valley. An unrelated increase in western Nevada (M5) consists of aftershocks of an $M = 5.3$ earthquake (1 August 1999). Rate increase near San Geronio pass (SG) southwest of the Hector Mine aftershock zone began a week before the mainshock, as did the increase north of Ventura (V), whereas the events near the Geysers (G) occurred two weeks after it. **b**, Landers response. Most remote triggering occurred to the north-northwest. The strong increase marked LSM is the (triggered) $M = 5.6$ Little Skull Mountain earthquake and its aftershocks.

dynamic triggering thresholds. These thresholds vary from a few tenths to a few MPa in most places, depend on local conditions, and exceed inferred static thresholds by more than an order of magnitude. At some sites, the onset of triggering was delayed until after the dynamic deformations subsided. Physical mechanisms consistent with all these observations may be similar to those that give rise to liquefaction or cyclic fatigue.

We have characterized the spatial seismicity rate increase associated with the Hector Mine and Landers earthquakes using maps of a smoothed β -statistic, which is sensitive to a contrast of average seismicity rate between two time intervals⁷ (Fig. 1). These maps show that after the unilaterally northward-rupturing Landers earthquake, seismicity increased to the north^{1,8}. Although the Hector Mine earthquake rupture seems to have been more complex than that of the Landers rupture, southward focusing of the radiated seismic energy or 'directivity' is suggested by relatively greater slip and fault length south of the hypocentre, and a slow rupture velocity^{9,10}. After the Hector Mine earthquake, remote (that is, beyond the aftershock zone or about one fault length) increases in the seismicity rate occurred to the south in three clusters; near Indio, the Salton Sea and Cerro Prieto. Rate increases were also observed after (and to the north) of both earthquakes at Long Valley¹¹. The post-Hector Mine seismicity increase at Long Valley consisted nearly entirely of magnitude (M) < 2.0 events. Time histories of cumulative seismicity (Fig. 2) verify the post-Hector Mine rate increases in these areas. The short or non-existent delays (intervals between mainshock and first recorded earthquake) suggest a causal link and provide insight for differentiating between physical models of triggering^{12,13}. For example, although triggering models invoking Coulomb failure, frictional instabilities, or sub-critical crack propagation predict delayed failure of an individual

fault, they do not predict delayed rate changes^{12–15}. Moreover, when the causative deformation is dynamic, these models predict that the rate changes last only as long as the driving transient¹³.

The first catalogued post-Hector Mine events occurred 36 min, 1 h, 1.35 days, and 40 mins after the mainshock at Indio, Salton Sea, Cerro Prieto, and Long Valley, respectively. Continuously recorded ground motions show a vigorous increase in local micro-earthquake activity near Salton Sea that started during the passage of the Hector Mine waves; no increase at Indio before the first catalogued earthquake; and a delay of about 10 min at Long Valley. The delays at Indio and Salton Sea are consistent with immediate triggering, given the average post-Hector Mine inter-event times¹. Rate increases at Cerro Prieto and Long Valley were not immediate; however, identification of early events can be difficult and no continuous data exist for Cerro Prieto. At all sites the rate increases last longer than the duration of dynamic deformations.

The static stress change represented as the Coulomb failure function (ΔCFF)¹⁶ and its relationship to seismicity rate changes have been considered as a triggering mechanism^{14–20}. We conclude that static stress changes alone could not have caused the remote seismicity rate changes after the Hector mine earthquake. We calculated ΔCFF for the Hector Mine earthquake on optimally oriented and other plausibly oriented planes. In the vicinity of Salton Sea, Cerro Prieto and Long Valley, the maximum absolute value of ΔCFF does not exceed 0.002 MPa, which is smaller than generally inferred static ΔCFF triggering thresholds (about 0.01 MPa)^{16,19}. Moreover, ΔCFF is nearly spatially symmetrical, whereas the more remote seismicity increases occurred primarily to the south of the rupture.

The correlation between the directions of rupture and seismicity rate changes for both the Landers and Hector Mine earthquakes

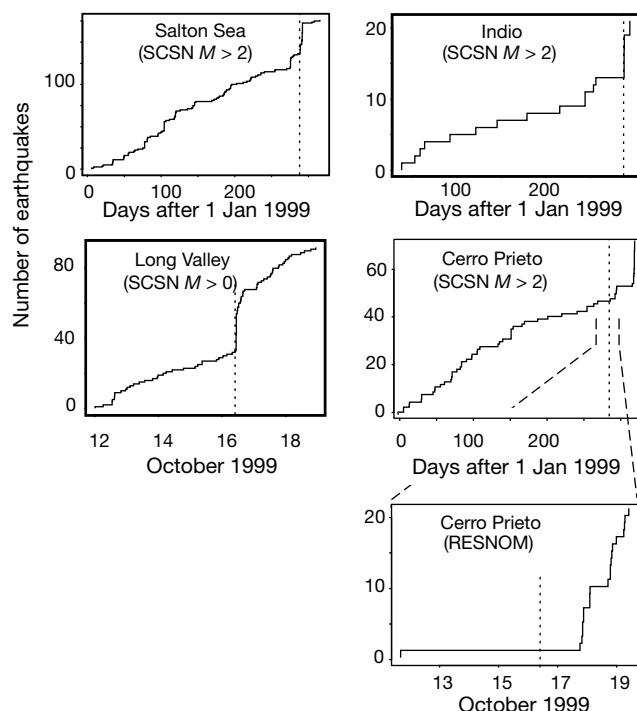


Figure 2 Time histories of seismicity for the areas showing significantly increased rates after the Hector Mine earthquake. Cumulative number of earthquakes in circles of 25 km radius centred on the south end of the Salton Sea, Indio, Long Valley and Cerro Prieto during indicated periods, listed in the SCSN and NCSN. RESNOM (Red Sismica del Noroeste de Mexico) catalogue data from near Cerro Prieto show a delayed rate increase, which was not observed elsewhere for $M > 2$ earthquakes. RESNOM is operated by the Centro de Investigacion Cientifica y Educacion Superior de Ensenada, Baja California, Mexico. Dashed lines indicate the time of the Hector Mine earthquake.

suggests a relationship between radiated seismic deformations and triggered seismicity. Recorded seismic ground motion velocities superimposed on the β -statistic maps corroborates this correlation (Fig. 3). Seismic wave velocities are approximately proportional to dynamic strains, both theoretically and empirically²¹. Dynamic strains constrain stresses, assuming the two are linearly related by means of typical crustal elastic moduli. The southern narrow band of seismicity increase after Hector Mine coincides in azimuth with the large, pulse-like ground velocities that probably result, at least in part, from the southward rupture directivity (Fig. 3a). A similar coincidence is observed for the Landers earthquake, but in the northward direction (Fig. 3b). Some of the large amplitudes south of both ruptures undoubtedly reflect site amplification due to thick Imperial Valley sediments²². However, the effects of directivity are still indicated in the Hector Mine waveforms, which are more pulse-like and larger to the south than those for the Landers earthquake, which had a larger magnitude.

The proximity and similar source mechanisms of the Landers²³ and Hector Mine^{9,10} earthquakes permit useful comparisons of their associated deformation fields because the point-source radiation pattern, propagation, and site effects should be approximately the same, and thus effectively cancel. Because direct measurements cannot be made at seismogenic depths, we assume that the relative magnitudes of surface deformations for both earthquakes are about the same as those at greater depth.

We propose two processes that may explain triggering by oscillatory deformations. Faults may be weakened by increased pore pressures associated with the compaction of saturated fault gouge (as in liquefaction) or by cyclic fatigue. These processes require deformations to exceed some threshold value. Measurements of peak velocities combined with estimates of seismicity rate change associated with each site allow us to test for the existence of a dynamic triggering threshold; to consider whether this threshold varies from site to site; and to estimate approximate bounds on its

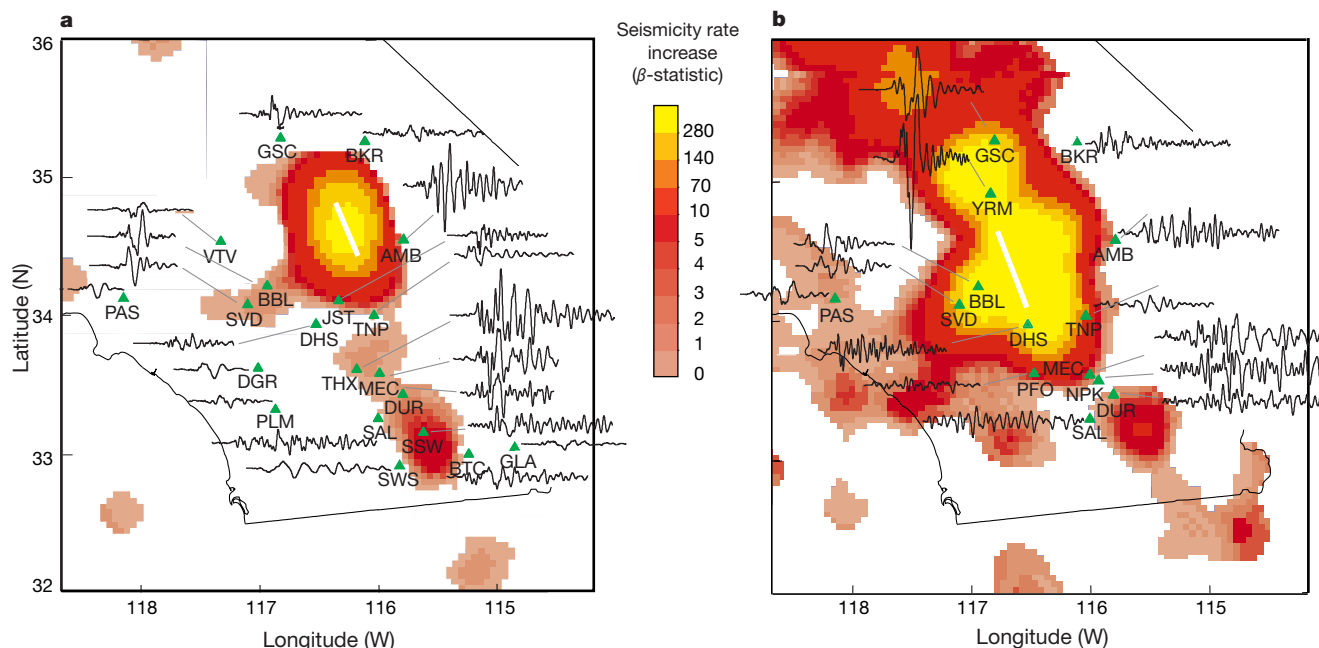


Figure 3 Comparison of spatial patterns of seismicity rate increases and dynamic deformation. **a, b**, Recorded horizontal seismic velocities serve as proxies for the dynamic deformation for the Hector Mine (**a**) and Landers (**b**) earthquakes, and are superimposed on portions of the β -statistic maps of Fig. 1. Velocity recordings resulted from rotating horizontal components to the direction yielding the maximum absolute value of velocity. The vertical component seismograms, which are much smaller at all stations, are not shown. Waveforms are plotted at the same vertical and horizontal scales. Waveform data

are derived from accelerations recorded at strong motion stations operated by the California Division of Mines and Geology, or are part of TriNet. Accelerations were filtered in the passband 0.05–0.5 Hz and integrated to velocity. Data were selected from all stations that recorded both of the earthquakes on scale and others to provide azimuthal coverage. Diagonal white lines indicate surface traces of the Hector Mine and Landers ruptures.

magnitude. If a dynamic triggering threshold exists, then at sites where seismicity increases occurred only after the Hector Mine earthquake, Hector Mine peak strains (velocities) would exceed those from Landers. An analogous relationship would be expected at sites only triggered by the Landers earthquake. Recorded velocities for both earthquakes at the same sites (Fig. 4a), albeit scant, are consistent with this hypothesis at both remote and nearer distances, suggesting that both static and dynamic deformations may encourage triggering. At the five Landers-only triggered sites, Landers peak velocities exceed those from Hector Mine at the two remote sites (GSC, PAS) but only at one (DHS) of the three closer stations. Of the six sites triggered (or possibly triggered) by the Hector Mine earthquake alone, both mainshocks were recorded at only two stations and Hector Mine peak velocities exceeded those from Landers at both of these (Fig. 4a).

Accepting the existence of a dynamic triggering threshold, we also can estimate bounds on its value(s) and examine its spatial variability. When both earthquakes are recorded at the same site, peak velocities provide both upper and lower bounds on a triggering threshold (Fig. 4a). When only one earthquake is recorded at a site, the peak velocity provides either an upper or lower bound (Fig. 4b and Supplementary Information). Although a single threshold value cannot satisfy the bounds at all sites, the observations suggest that thresholds span a small range from between a few to a few tens of cm s^{-1} (10 cm s^{-1} corresponds to roughly 1 MPa and 30

microstrain, assuming a rigidity of 3×10^4 MPa and a constant horizontal wave velocity of 3 km s^{-1}), whereas the measurements at Long Valley suggest that the triggering threshold must be less than 1 cm s^{-1} (0.1 MPa or 3 microstrain). The reason for this marked sensitivity at Long Valley is a subject for further research. All these threshold values are only suggestive, as the data are at the surface; we do not resolve the deformations onto fault planes, and deformations at greater depth undoubtedly differ (probably by less than a factor of four^{3,4}).

The 1999, $M_w = 7.1$ Hector Mine earthquake provided a rare repeat experiment on dynamic triggering. The correlation between the direction of rupture, of enhanced seismic ground velocities, and increased seismicity rate for this and the Landers earthquake provides strong evidence of a causal relationship. The proximity and similarity of these two earthquakes allowed us to compare the observed deformation fields directly, thereby providing additional support, independent of model calculations, that the dynamic deformations were probably triggering agents. It also allowed us to test for the existence of and to estimate the magnitudes of dynamic triggering thresholds. These estimates are generally consistent with peak stresses inferred at sites of seismicity rate increase after the Landers earthquake^{1–5}, and exceed static ΔCFF triggering thresholds by more than an order of magnitude. Unlike static stress changes, transient deformations cannot permanently alter the applied stress (for example, that due to long-term plate motion)

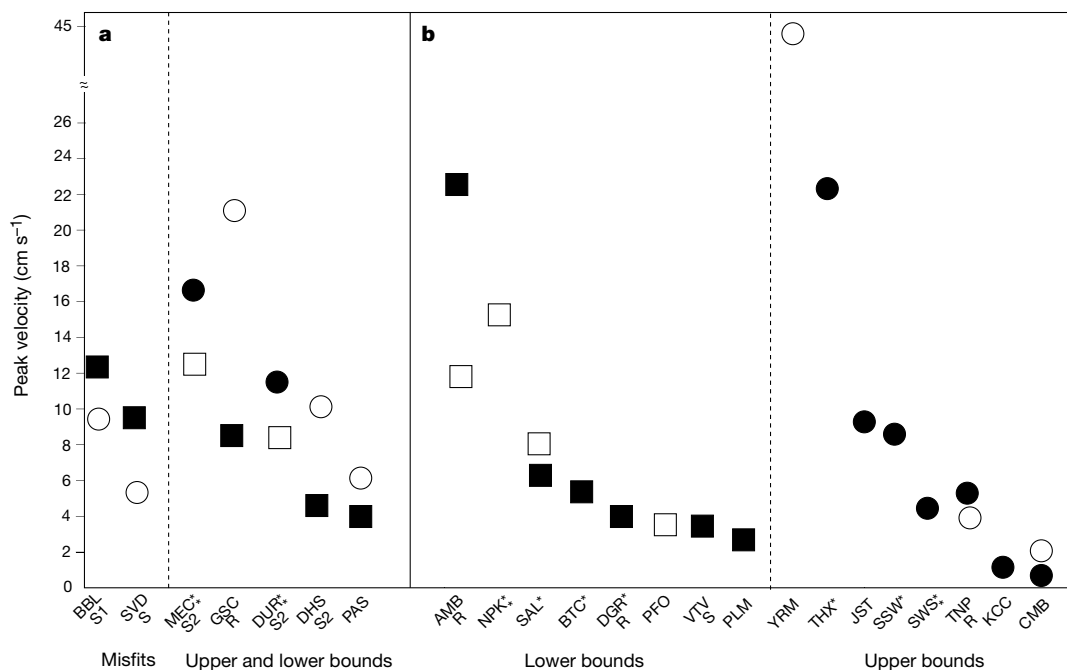


Figure 4 Approximate bounds on dynamic triggering velocity thresholds, which serve as proxies for triggering stress or strain thresholds. Data points indicate the maximum velocities recorded at each site (ordinate values) for the Hector Mine and/or Landers earthquakes. Circles or squares indicate triggered seismicity was or was not observed, respectively. Filled symbols correspond to the Hector Mine event. **a**, Peak velocities at sites that recorded both mainshocks and at which only one mainshock triggered a rate change. If a dynamic triggering threshold exists, the velocity from the triggering mainshock must exceed that from the non-triggering mainshock. Measurements at stations SVD and BBL are inconsistent with this hypothesis, suggesting that the triggering process at these sites involves more than exceeding a peak dynamic triggering threshold. The consistent observations provide upper and lower bounds on a triggering threshold value, and indicate that thresholds must depend on local site conditions; for example, threshold bounds at MEC and DUR do not overlap. **b**, Sites recording only one peak velocity. If a threshold exists, then at sites where all the velocities recorded are for mainshocks that triggered seismicity (one or two circles), the smaller peak velocity

provides an upper bound. At sites where all the velocities recorded are for mainshocks that did not trigger seismicity (one or two squares), the larger peak velocity provides a lower bound. These observations also indicate that thresholds must depend on local site conditions, and all fall in a range between a few and a few tens of cm s^{-1} . Peak velocities are measured from rotated horizontal components (Fig. 3). A site is associated with triggering when the seismometer is less than $\sim 5 \text{ km}$ from a region where the β -statistic exceeds two and time histories show abrupt seismicity increase (see Fig. 2). This distance roughly corresponds to a quarter-wavelength of the largest amplitude waves and thus the approximate distance within which dynamic deformations are nearly constant¹⁹. At the station SSW the seismicity increase began about one week after the Landers earthquake. Asterisks, stations that recorded ground motions continuously; double asterisks, sites where the seismicity increase after Hector Mine was not evident but the sites are along strike and between stations of clear triggering (Fig. 3). Local site conditions: R, rock; S1, stiff sediments $< 60 \text{ m}$ thick; S2, stiff sediments $> 60 \text{ m}$ thick; S, soil of unknown thickness. CMB and KCC are the two stations closest to Long Valley.

and thus can only enhance the likelihood of failure by altering some physical property of, or near, the triggered fault⁸. Our hypotheses and theoretical calculations²⁴ suggest that such alteration requires large deformations of the order of the dynamic thresholds estimated here. Although our observations indicate that transient dynamic deformations trigger earthquakes, the mechanism(s) by which they do so remain unknown^{11,13,25–28}. Evidence presented here for a possible triggering threshold and for significant time delays (in some cases) between the shaking and seismicity rate increases provide insights for elucidating the triggering mechanism. □

Received 13 April 2000; accepted 23 March 2001.

- Hill, D. P. *et al.* Remote seismicity triggered by the M7.5 Landers, California earthquake of June 28, 1992. *Science* **260**, 1617–1623 (1993).
- Anderson, J. G. *et al.* Seismicity in the western Great Basin apparently triggered by the Landers, California, earthquake 28 June 1992. *Bull. Seismol. Soc. Am.* **84**, 863–891 (1994).
- Gomberg, J. & Bodin, P. Triggering of the Little Skull Mountain, Nevada earthquake with dynamic strains. *Bull. Seismol. Soc. Am.* **84**, 844–853 (1994).
- Spudich, P., Steck, L. K., Hellweg, M., Fletcher, J. B. & Baker, L. M. Transient stresses at Parkfield, California, produced by the M 7.4 Landers earthquake of June 28, 1992: observations from the UPSAR dense seismograph array. *J. Geophys. Res.* **100**, 675–690 (1995).
- Gomberg, J. Stress/strain changes and triggered seismicity following the Ms7.4 Landers, California earthquake. *J. Geophys. Res.* **101**, 751–764 (1996).
- Gomberg, J. & Davis, S. Stress/strain changes and triggered seismicity at The Geysers, California. *J. Geophys. Res.* **101**, 733–749 (1996).
- Matthews, M. V. & Reasenberg, P. A. Statistical methods for investigating quiescence and other temporal seismicity patterns. *Phys. Appl. Geophys.* **126**, 357–372 (1988).
- Kilb, D., Gomberg, J. & Bodin, P. Earthquake triggering by dynamic stresses. *Nature* **408**, 570–574 (2000).
- Dreger, D. & Kaverina, A. Seismic remote sensing for the earthquake source process and near-source strong shaking: a case study of the October 16, 1999 Hector Mine earthquake. *Geophys. Res. Lett.* **27**, 1941–1944 (2000).
- Ji, C., Wald, D. J. & Helmberger, D. V. Source description of the 1999 Hector Mine, California earthquake; Part II: complexity of slip history. *Bull. Seismol. Soc. Am.* (in the press).
- Hill, D. P., Johnston, M. J. S., Langbein, J. O. & Bilham, R. Response of Long Valley caldera to the Mw = 7.3 Landers, California, earthquake. *J. Geophys. Res.* **100**, 12985–13005 (1995).
- Gomberg, J., Beeler, N. & Blanpied, M. On rate-state and Coulomb failure models. *J. Geophys. Res.* **105**, 7857–7871 (2000).
- Gomberg, J. The failure of earthquake failure models. *J. Geophys. Res.* (in the press).
- Das, S. & Scholz, C. H. Theory of time-dependent rupture in the Earth. *J. Geophys. Res.* **86**, 6039–6051 (1981).
- Dieterich, J. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.* **99**, 2601–2618 (1994).
- Harris, R. A. Introduction to special section: stress triggers, stress shadows, and implications for seismic hazard. *J. Geophys. Res.* **103**, 24347–24358 (1998).
- King, G. C. P., Stein, R. S. & Lin, J. Static stress changes and the triggering of earthquakes. *Bull. Seismol. Soc. Am.* **84**, 935–953 (1994).
- Hardebeck, J. L., Nazareth, J. J. & Hauksson, E. The static stress change triggering model: constraints from two southern California aftershock sequences. *J. Geophys. Res.* **103**, 24427–24438 (1998).
- Reasenberg, P. A. & Simpson, R. W. Response of regional seismicity to the static stress change produced by the Loma Prieta earthquake. *Science* **255**, 1687–1690 (1992).
- Ziv, A. & Rubin, A. M. Static stress transfer and earthquake triggering: no lower threshold in sight? *J. Geophys. Res.* **103**, 13631–13642 (2000).
- Gomberg, J. & Agnew, D. C. The accuracy of seismic estimates of dynamic strains from Pinon Flat Observatory, California, strainmeter and seismograph data. *Bull. Seismol. Soc. Am.* **86**, 212–220 (1996).
- Cramer, C. H. & Darragh, R. B. Peak accelerations from the 1992 Landers and Big Bear, California earthquakes. *Bull. Seismol. Soc. Am.* **84**, 589–595 (1994).
- Wald, D. J. & Heaton, T. H. Spatial and temporal distribution of slip for the 1992 Landers, California earthquake. *Bull. Seismol. Soc. Am.* **84**, 668–691 (1994).
- Gomberg, J., Beeler, N. M., Blanpied, M. L. & Bodin, P. Earthquake triggering by transient and static deformations. *J. Geophys. Res.* **103**, 24411–24426 (1998).
- Nur, A. & Booker, J. R. Aftershocks caused by pore fluid flow? *Science* **175**, 885–887 (1972).
- Linde, A., Sacks, I., Johnston, M., Hill, D. & Bilham, R. Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity. *Nature* **371**, 408–410 (1994).
- Sturtevant, B., Kanamori, H. & Brodsky, E. E. Seismic triggering by rectified diffusion in geothermal systems. *J. Geophys. Res.* **101**, 25269–25282 (1996).
- Sornette, D. & Sornette, C. G. Complex critical exponents from renormalization group theory of earthquakes: implications for earthquake predictions. *J. Phys. Ser. I* **5**, 607–619 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

Seismic data were obtained from the Northern California Earthquake Data Center operated by the Berkeley Seismological Laboratory, University of California, Berkeley, and from the Southern California Seismic Network and TriNet operated by the US Geological Survey, the California Institute of Technology, and the California Division of Mines and Geology. R. Castro and C. Rebolgar of CICESE also provided data. We also thank D. Hill, L. Jones, C. Marone and D. Agnew for comments on the manuscript.

Correspondence and requests for materials should be addressed to J.G. (e-mail: gomberg@usgs.gov).

Limited carbon storage in soil and litter of experimental forest plots under increased atmospheric CO₂

William H. Schlesinger* & John Lichter†

* Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina 27708, USA

† Biology Department and Environmental Studies Program, Bowdoin College, Brunswick, Maine 04011, USA

The current rise in atmospheric CO₂ concentration is thought to be mitigated in part by carbon sequestration within forest ecosystems^{1,2}, where carbon can be stored in vegetation or soils. The storage of carbon in soils is determined by the fraction that is sequestered in persistent organic materials, such as humus. In experimental forest plots of loblolly pine (*Pinus taeda*) exposed to high CO₂ concentrations^{3,4}, nearly half of the carbon uptake is allocated to short-lived tissues, largely foliage. These tissues fall to the ground and decompose, normally contributing only a small portion of their carbon content to refractory soil humic materials⁵. Such findings call into question the role of soils as long-term carbon sinks, and show the need for a better understanding of carbon cycling in forest soils. Here we report a significant accumulation of carbon in the litter layer of experimental forest plots after three years of growth at increased CO₂ concentrations (565 μl l⁻¹). But fast turnover times of organic carbon in the litter layer (of about three years) appear to constrain the potential size of this carbon sink. Given the observation that carbon accumulation in the deeper mineral soil layers was absent, we suggest that significant, long-term net carbon sequestration in forest soils is unlikely.

The FACE (free air CO₂ enrichment) experiment in the Duke Forest (Orange County, North Carolina, USA) is composed of six plots 30-m in diameter^{3,4}. Three experimental plots are fumigated with CO₂ to maintain the atmospheric CO₂ concentration 200 μl l⁻¹ above ambient. Three control plots are fumigated with ambient air only (365 μl l⁻¹). The experiment began on 27 August 1996 and is continuous (24 h d⁻¹; 365 d yr⁻¹) except in extreme weather⁴. During the first two years of the experiment, net primary production increased by 25% in the fumigated plots³.

The forest is derived from three-year-old pine seedlings that were planted in 1983 in a 2.4 × 2.4 m spacing. In 1996, the forest was approximately 14 m tall, with a closed canopy and pine accounting for 98% of the basal area. The 32-ha site contains an elevation gradient of 15 m between the highest and lowest points, but topographic relief is ≤1° throughout. Soils are of the Enon Series, a low-fertility Ultic Alfisol, which is typical of many upland areas of the southeastern United States. The soil is derived from igneous rock, yielding a relatively acidic (pH = 5.75), well-developed soil profile with mixed clay mineralogy. Boreholes show up to 1 m of topsoil, underlain by 5 m of saprolite, with highly fractured granodiorite to diorite bedrock at greater depth. The static water table is at 6 m, but the site is poorly drained from late autumn until early spring, leading to saturated surface soils. The mean annual temperature is 15.5 °C and mean annual precipitation is 1,140 mm.

The CO₂ used for fumigation is derived from natural gas and strongly depleted in ¹³C: δ¹³C = -43.0 ± 0.6 (mean ± s.e.; in ‰) versus the Pee Dee Belemnite standard, PDB. Using this CO₂ to increase the atmospheric concentration by 200 μl l⁻¹ changes the δ¹³C of atmospheric CO₂ in the FACE plots from -8‰ to -20‰. As a result of the photosynthetic fractionation, needles grown under increased CO₂ have a δ¹³C value of -39.3‰ (ref. 6), and the δ¹³C

value of new roots less than 2 mm in diameter ranges from -38.5% to -39% (R. Matamala, unpublished results). We follow the incorporation of this isotopic signal into soil organic components as an index of their dynamics.

During construction of the experiment, pretreatment soil samples were collected and archived from 16 excavations surrounding each plot. In early October 1999, after three growing seasons of fumigation, twelve soil cores 4.76 cm in diameter were collected from stratified, random positions within each plot. Each sample consisted of the undecomposed plant materials on the soil surface (that is, the forest floor) and soil layers extending from 0 to 15 cm and from 15 to 30 cm depths, which collectively include nearly all of the root biomass⁷. The leaves fall in late October, so this sampling time corresponded to the annual minimum in forest-floor mass. The volume and mass of stones were determined for each sample and subtracted from the sample mass to estimate soil bulk density (g cm^{-3}). All samples were dried at 48°C for 5 days, and mineral soils were sieved ($<2\text{ mm}$) to remove coarse roots and stones.

The organic material in 30-g subsamples of the mineral soil was separated into coarse particulate organic matter (POM; $>0.5\text{ mm}$), fine POM ($53\text{ }\mu\text{m}$ – 0.5 mm), and mineral-associated ($<53\text{ }\mu\text{m}$) fractions^{8,9}. During decomposition, fresh plant litter is transformed sequentially into these fractions, which have increasing stability and potential to account for long-term carbon sequestration in soils¹⁰. The fraction $<53\text{ }\mu\text{m}$, including dissolved organic matter, was obtained following centrifugation (2,000 r.p.m. for 10 min). Samples of each fraction and bulk samples of the soil and forest floor were analysed for %C, %N, and $\delta^{13}\text{C}$ using a SIRA Series II isotope ratio mass spectrometer (Micromass, Manchester, UK), operated in automatic trapping mode after Dumas combustion of samples in a C and N analyser (NA1500 Series 1, Carlo Erba Instrumentazione, Milan, Italy). The reference CO_2 , calibrated against the PDB standard, was obtained from Oztech (Dallas, Texas). Data are expressed as $\delta^{13}\text{C}$ after correction for the isotope contribution of the O_2 used in sample combustion.

The mean of samples within each plot is considered to be the experimental unit for statistical analyses, giving a sample size of three each for the experimental and control plots. There were no

significant differences in the mass of the forest floor between control and experimental plots at the beginning of the experiment¹¹. In pretreatment samples of soil from 0–15 cm depth, %C was 1.432% in control plots and 1.542% in fumigated plots, but these mean values were also not significantly different ($F = 0.388$, $P = 0.567$, $n = 3$ in one-way analysis of variance, ANOVA). In subsequent evaluations of differences in forest floor and mineral soil parameters between control and fumigated plots, we used a one-way ANOVA, incorporating a covariate to account for initial conditions.

At the end of three years of CO_2 fumigation, significant differences were seen between control and fumigated plots for total mass and the carbon and nitrogen content of the forest floor, %C in the mineral soil at 0–15 cm depth, and ^{13}C in the forest floor and in bulk and size-fractionated carbon in the mineral soil from 0–15 cm depth (Table 1). No significant differences were seen in soils below 15 cm depth. Differences in the forest floor derive solely from an increase in its mass in CO_2 -fumigated plots, because %C and %N were not altered by the experiment. Greater %C, without a concomitant change in %N, contributed to a greater pool of carbon in the mineral soil of fumigated plots, but this difference in C mass was not statistically significant. A two-way ANOVA, with time (1996 versus 1999) and treatment (365 and $565\text{ }\mu\text{L l}^{-1}\text{ CO}_2$) as main effects, indicates no significant effect of time ($P = 0.70$) or treatment ($P = 0.085$) on %C in the 0–15 cm soil depth of control or fumigated plots nor any interaction between time and treatment ($P = 0.41$). This result indicates that the difference in %C in the 0–15 cm soil layer is related to initial plot differences and not to the experimental treatment.

In non-steady-state conditions, the mean residence time (MRT) of carbon pools can be estimated from: $C_t = C_0 e^{-kt} + (I/k)(1 - e^{-kt})$ solved by iteration for k , where C_t is the carbon stock after three years of fumigation, C_0 is the pre-fumigation stock, t is the duration of the experiment, I is the annual C input from litterfall and root turnover, and k is the decomposition constant¹². For the control and fumigated plots, the MRT values ($=1/k$) for C in the forest floor were 2.50 and 2.86 yr, respectively.

Differences in $\delta^{13}\text{C}$ result for the entry of new plant debris

Table 1 Forest floor and soil parameters (mean $\pm$ s.e.) after three years of CO_2 fumigation

Forest floor									
	Mass (g m ⁻²)	%C	%N	C:N	Total C (g m ⁻²)	Total N (g m ⁻²)	δ ¹³ C		
Control plots	1,833.5 (206.5)	38.24 (0.61)	0.90 (0.02)	42.71 (1.06)	700.98 (87.23)	16.55 (2.24)	-28.90 (0.27)		
Fumigated plots	2,359.6 (32.4)	37.68 (0.41)	0.88 (0.02)	43.36 (0.74)	883.92 (24.30)	20.57 (0.75)	-34.15 (0.71)		
<i>P</i>	0.008	0.484	0.471	0.637	0.014	0.009	0.002		
Mineral soil 0–15 cm depth									
	%C	%N	C:N	Total C (g m ⁻²)	Total N (g m ⁻²)	δ ¹³ C (‰)			
						Bulk soil	Coarse POM	Fine POM	<53 μm fraction
Control plots	1.31 (0.07)	0.07 (0.01)	18.09 (0.91)	1,901.43 (51.63)	105.8 (8.5)	-26.244 (0.182)	-26.819 (0.272)	-27.144 (0.120)	-25.831 (0.199)
Fumigated plots	1.59 (0.07)	0.08 (0.01)	18.98 (1.18)	2,207.65 (136.33)	118.6 (13.6)	-28.054 (0.195)	-29.746 (0.106)	-28.479 (0.078)	-27.214 (0.127)
<i>P</i>	0.037	0.839	0.581	0.180	0.570	0.002	0.001	0.001	0.004
Mineral soil 15–30 cm depth									
	%C	%N	C:N	Total C (g m ⁻²)	Total N (g m ⁻²)	δ ¹³ C (‰)			
						Bulk soil	Coarse POM	Fine POM	<53 μm fraction
Control plots	0.484 (0.054)	0.030 (0.003)	16.1 (0.47)	657.1 (61.3)	40.8 (3.9)	-24.199 (0.457)	-24.334 (0.381)	-26.110 (0.261)	-24.096 (0.507)
Fumigated plots	0.539 (0.057)	0.033 (0.004)	16.2 (0.52)	747.4 (46.4)	46.2 (1.7)	-24.955 (0.554)	-25.451 (0.156)	-26.840 (0.268)	-24.515 (0.585)
<i>P</i>	0.606	0.356	0.946	0.263	0.286	0.352	0.053	0.123	0.617

The *P* value is derived from one-way ANOVA used to test for treatment effects, with a covariate included in the model to account for initial site differences in forest floor mass and %C and %N of the mineral soil layers. One collection of forest floor mass and one measurement of %C in the mineral soil at 0–15 cm depth were deleted from the analysis following a test for outliers³⁰. POM, particulate organic matter.

into forest floor materials and into coarse POM of the mineral soil. In the fumigated plots, more than 50% of the organic carbon in the forest floor is derived from plant tissues grown since the beginning of the experiment. The change in $\delta^{13}\text{C}$ in forest floor carbon can be used to provide an alternative estimate of its MRT in the fumigated plots. We assume an exponential decline in $\delta^{13}\text{C}$ in plant litter after the onset of fumigation, and derive k by iteration using the formula: $1 - f = e^{-kt}$, where f is the fraction of organic matter replaced by new carbon with depleted $\delta^{13}\text{C}$ (that is, $f = (\delta^{13}\text{C}_{\text{increased}} - \delta^{13}\text{C}_{\text{ambient}})/(\delta^{13}\text{C}_{\text{new}} - \delta^{13}\text{C}_{\text{ambient}})$) (ref. 13). Assuming that $\delta^{13}\text{C}$ of fresh litter (C_{new}) in the fumigated plots is -37.5‰ (ref. 14), the MRT for carbon in the forest floor is 3.23 yr ($k = 0.31$)—in good agreement with the literature¹⁵ and with the estimate derived above ($k = 0.35$).

Changes in $\delta^{13}\text{C}$ in bulk soil organic matter at 0–15 cm depth yield an estimated 20-year mean residence time for this pool. The greatest change in $\delta^{13}\text{C}$ is seen in coarse POM and the least in the mineral-associated fraction, consistent with the increasing stability of carbon in finer fractions of soil organic matter^{8–10}. A small change in $\delta^{13}\text{C}$ in the coarse POM at 15–30 cm depth indicates inputs of fresh plant debris at that depth, but the absence of change in $\delta^{13}\text{C}$ in the mineral-associated fraction suggests that there has been little production or downward movement of dissolved organic carbon compounds, such as fulvic acids, during the three-year period.

Growth of this forest under increased CO_2 has increased the production of dead plant debris, which is delivered to the forest floor and to the mineral soil through plant litterfall^{3,14}, leaching of above-ground plant components by rain¹⁶, and root turnover⁷. The cumulative change in these inputs, owing to growth at high CO_2 , is about 166 g C m^{-2} for the three growing seasons of CO_2 fumigation (Table 2). This higher input of organic materials can—within experimental error (Table 1)—account for the greater mass of carbon held in the forest floor of CO_2 -fumigated plots at the end of the third growing season (183 g C m^{-2}). Additional carbon may derive from root exudations of low-molecular-weight carbon compounds (for example, see ref. 17), which are reported to increase when plants are grown at high CO_2 (refs 18, 19). The non-significant change in the mass of organic carbon in the mineral soil is similar to the results of several CO_2 -fumigation experiments in grasslands, which demonstrate substantial turnover of soil carbon pools with little or no net accumulation in the profile^{20,21}. These results are in contrast to recent predictions of a large potential sink for C in soils^{22,23}.

Although soil carbon accumulates rapidly during the initial growth of pine plantations in the southeastern United States²⁴, wood growth is normally the largest sink for carbon in these forests^{24,25}. In the Duke forest experiment, trees in the control plots sequestered about $430 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 3) during their first 16 years of growth, and the forest floor has accumulated at about $44 \text{ g C m}^{-2} \text{ yr}^{-1}$, similar to rates measured in 15-year old plantations in Virginia²⁵. Such reforested lands in the eastern United States sequester 0.07 to $0.17 \text{ Pg C yr}^{-1}$ (refs 26, 27)—about 10% of the US

fossil-fuel emissions, currently estimated at 1.4 Pg C yr^{-1} (ref. 28). Of the total estimated US sink of $0.15\text{--}0.35 \text{ Pg C yr}^{-1}$ in vegetation and soils², Schimel *et al.*²⁸ suggest that only $0.08 \text{ Pg C yr}^{-1}$, or less than $20 \text{ g C m}^{-2} \text{ yr}^{-1}$, has occurred in forests as a result of CO_2 fertilization of plant growth during the past century—changes in land use account for a larger fraction²⁹.

In response to the step-function increase of $200 \mu\text{l l}^{-1}$ in atmospheric CO_2 , this southern pine forest shows an incremental sink for C (183 g C m^{-2}) in the forest floor at the end of three years of fumigation (Table 2). The decomposition constant (k) estimated for the fumigated plots indicates that at equilibrium (C_{∞}), the forest floor under increased CO_2 will contain $1,103 \text{ g C m}^{-2}$ ($C_{\infty} = I/k$ and $k = 0.31$). Thus, 80% of the equilibrium forest-floor mass, and 46% of the incremental sequestration due to CO_2 fertilization, has already accumulated. We expect that the incremental sequestration due to growth at high CO_2 will be substantially less than $60 \text{ g C m}^{-2} \text{ yr}^{-1}$ in future years. In the absence of significant changes in the carbon content of the mineral soil, these data suggest only a limited potential for long-term soil carbon sequestration in this forest in response to rising CO_2 in Earth's atmosphere. The larger sink for carbon in vegetation, compared to soils, differs from the results of recent models of the response of the biosphere to climate change and rising CO_2 (ref. 23). □

Received 3 January; accepted 2 April 2001.

- Ciais, P., Peylin, P. & Bousquet, P. Regional biospheric fluxes as inferred from atmospheric CO_2 measurements. *Ecol. Appl.* **10**, 1574–1589 (2000).
- Houghton, R. A., Hackler, J. L. & Lawrence, K. T. The U.S. carbon budget: contributions from land-use change. *Science* **285**, 574–578 (1999).
- DeLucia, E. H. *et al.* Net primary production of a forest ecosystem with experimental CO_2 enrichment. *Science* **284**, 1177–1179 (1999).
- Hendrey, G. H., Ellsworth, D. S., Lewin, K. F. & Nagy, J. A free-air enrichment system for exposing tall forest vegetation to elevated atmospheric CO_2 . *Glob. Change Biol.* **5**, 293–309 (1999).
- Schlesinger, W. H. Evidence from chronosequence studies for a low carbon-storage potential of soils. *Nature* **348**, 232–234 (1990).
- Ellsworth, D. S. CO_2 enrichment in a maturing pine forest: Are CO_2 exchange and water status in the canopy affected? *Plant Cell Environ.* **22**, 461–472 (1999).
- Matamala, R. & Schlesinger, W. H. Effects of elevated atmospheric CO_2 on fine root production and activity in an intact temperate forest ecosystem. *Glob. Change Biol.* **6**, 967–980 (2000).
- Cambardella, C. A. & Elliott, E. T. Methods for physical separation and characterization of soil organic matter fractions. *Geoderma* **56**, 449–457 (1993).
- Gill, R. A., Burke, I. C., Lauenroth, W. K. & Milchunas, D. G. Relationship between root biomass and soil organic matter pools in the shortgrass steppe of eastern Colorado. *Ecosystems* **2**, 226–236 (1999).
- Jastrow, J. D., Boutton, T. W. & Miller, R. M. Carbon dynamics of aggregate-associated organic matter estimated by carbon-13 natural abundance. *Soil Sci. Soc. Am. J.* **60**, 801–807 (1996).
- Andrews, J. A. *Changes to Belowground Carbon Dioxide Dynamics under CO_2 Enrichment of a Forest Ecosystem*. Thesis, Duke Univ. (1999).
- Olson, J. S. Energy storage and the balance of producers and decomposers in ecological systems. *Ecology* **44**, 322–331 (1963).
- Balesdent, J., Wagner, G. H. & Mariotti, A. Soil organic matter turnover in long-term field experiments as revealed by carbon-13 natural abundance. *Soil Sci. Soc. Am. J.* **52**, 118–124 (1988).
- Finni, A. C., Allen, A. S., DeLucia, E. H., Ellsworth, D. S. & Schlesinger, W. H. Forest litter production, chemistry, and decomposition following two years of Free-Air CO_2 Enrichment. *Ecology* **82**, 470–484 (2001).
- Jorgensen, J. R., Wells, C. G. & Metz, L. J. Nutrient changes in decomposing loblolly pine forest floor. *Soil Sci. Soc. Am. J.* **44**, 1307–1314 (1980).
- Lichter, J., Lavine, M. L., Mace, K. A., Richter, D. D. & Schlesinger, W. H. Throughfall chemistry in a loblolly pine plantation under elevated atmospheric CO_2 concentrations. *Biogeochemistry* **50**, 73–93 (2000).
- Fox, T. R. & Comerford, N. B. Low-molecular-weight organic acids in selected forest soils of the southeastern USA. *Soil Sci. Soc. Am. J.* **54**, 1139–1144 (1990).
- DeLucia, E. H., Callaway, R. M., Thomas, E. M. & Schlesinger, W. H. Mechanisms of phosphorus acquisition for ponderosa pine under high CO_2 and temperature. *Ann. Bot.* **79**, 111–120 (1997).
- Jones, T. H. *et al.* Impacts of rising atmospheric carbon dioxide on model terrestrial ecosystems. *Science* **280**, 441–443 (1998).
- Hungate, B. A. *et al.* The fate of carbon in grasslands under carbon dioxide enrichment. *Nature* **388**, 576–579 (1997).
- van Kessel, C., Horwath, W. R., Hartwig, U., Harris, D. & Lüscher, A. Net soil carbon input under ambient and elevated CO_2 concentrations: isotopic evidence after 4 years. *Glob. Change Biol.* **6**, 435–444 (2000).
- Harrison, K., Broecker, W. & Bonani, G. A strategy for estimating the impact of CO_2 fertilization on soil carbon storage. *Glob. Biogeochem. Cycles* **7**, 69–80 (1993).
- Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled model. *Nature* **408**, 184–187 (2000).
- Richter, D. D., Markewitz, D., Trumbore, S. & Wells, C. G. Rapid accumulation and turnover of soil carbon in a re-establishing forest. *Nature* **400**, 56–58 (1999).
- Schiffman, P. N. & Johnson, W. C. Phytomass and detrital carbon storage during forest regrowth in the southeastern United States piedmont. *Can. J. Forest Res.* **19**, 69–78 (1989).

Table 2 Estimated inputs of organic matter to soils

	Control plots	Fumigated plots
Inputs		
Litterfall ¹⁴	744	883
Throughfall ¹⁶	24	31
Root turnover ⁷	37	57
Total inputs	805	971
Pools		
Forest floor	701	884
Soil 0–15 cm depth	1,901	2,208
Soil 15–30 cm depth	657	747
Total C pool	3,259	3,839

Inputs were measured from 1 January 1997 to 30 September 1999 (1,003 d, 2.75 yr), and soil carbon pools were measured in early October 1999 (compare Table 1).

26. Delcourt, H. R. & Harris, W. F. Carbon budget of the southeastern U.S. biota: analysis of historical change in trend from source to sink. *Science* **210**, 321–323 (1980).
27. Brown, S. L. & Schroeder, P. E. Spatial patterns of aboveground production and mortality of woody biomass for eastern U.S. forests. *Ecol. Appl.* **9**, 968–980 (1999).
28. Schimel, D. *et al.* Contribution of increasing CO₂ and climate to carbon storage by ecosystems in the United States. *Science* **287**, 2004–2006 (2000).
29. Casperson, J. P. *et al.* Contributions of land-use history to carbon accumulation in U.S. forests. *Science* **290**, 1148–1151 (2000).
30. Weisberg, S. *Applied Linear Regression* 114–117 (Wiley, New York, 1985).

Acknowledgements

We thank K. Harrison and J. Andrews for collecting pretreatment samples for covariance analysis; H. Hemric, R. Gill, M. Lavine and A. Mace for technical assistance; L. Giles for mass spectrometry analyses; and J. Clark, E. Davidson, L. Dellwo, A. Hirsch and D. Schimel for critical reviews of the manuscript. Operated in cooperation with Brookhaven National Laboratory, the Duke Forest FACE project is supported by the US Department of Energy.

Correspondence and requests for materials should be addressed to W.H.S. (e-mail: schlesin@duke.edu).

Soil fertility limits carbon sequestration by forest ecosystems in a CO₂-enriched atmosphere

Ram Oren*, David S. Ellsworth†‡, Kurt H. Johnsen§, Nathan Phillips||, Brent E. Ewers*, Chris Maier§, Karina V.R. Schäfer*, Heather McCarthy*, George Hendrey‡, Steven G. McNulty¶ & Gabriel G. Katul*

* Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina 27708, USA

† School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48109, USA

‡ Department of Environmental Sciences, Brookhaven National Laboratory, Upton, New York 11973-5000, USA

§ Southern Research Station, US Forest Service, Research Triangle Park, North Carolina 27709, USA

|| Department of Geography, Boston University, Boston, Massachusetts 02215, USA

¶ Southern Global Climate Change Program, US Forest Service, Raleigh, North Carolina 27606, USA

Northern mid-latitude forests are a large terrestrial carbon sink^{1–4}. Ignoring nutrient limitations, large increases in carbon sequestration from carbon dioxide (CO₂) fertilization are expected in these forests⁵. Yet, forests are usually relegated to sites of moderate to poor fertility, where tree growth is often limited by nutrient supply, in particular nitrogen^{6,7}. Here we present evidence that estimates of increases in carbon sequestration of forests, which is expected to partially compensate for increasing CO₂ in the atmosphere, are unduly optimistic⁸. In two forest experiments on maturing pines exposed to elevated atmospheric CO₂, the CO₂-induced biomass carbon increment without added nutrients was undetectable at a nutritionally poor site, and the stimulation at a nutritionally moderate site was transient, stabilizing at a marginal gain after three years. However, a large synergistic gain from higher CO₂ and nutrients was detected with nutrients added. This gain was even larger at the poor site (threefold higher than the expected additive effect) than at the moderate site (twofold higher). Thus, fertility can restrain the response of wood carbon sequestration to increased atmospheric CO₂. Assessment of future carbon sequestration should consider the limitations imposed by soil fertility, as well as interactions with nitrogen deposition.

By burning fossil fuel and forests, and converting land to intensive agriculture use, humans have elevated the atmospheric concentration of CO₂ (ref. 3) and the deposition of atmospheric nitrogen (N)^{9,10}. Growth of many tree species is enhanced with provisions of both CO₂ and N in the suboptimal range¹¹, making it difficult to assess the effect of increased availability of either one when the supply of both increases concurrently¹⁰. Experiments indicate decreased nutrient availability¹², owing to increased carbon (C)/N ratios in elevated CO₂-grown foliage and litter¹³, but increased nutrient uptake with elevated CO₂-induced growth enhancement of fine roots¹⁴. Because most forests occur on low-nutrient soils, the uncertain effects of elevated CO₂ on nutrient supply hinders our ability to estimate forest C sequestration for future global C budgets.

If tree nutrient uptake does not increase in proportion to the growth response to elevated CO₂, then maturing pine trees should show less elevated CO₂-induced growth enhancements on low-fertility sites than on moderate sites, and thus should respond more to elevated CO₂ under improved nutrition. To evaluate these predictions, we used two field experiments with related loblolly pine (*Pinus taeda* L.) genotypes: the longest running forest-based free-air CO₂ enrichment (FACE) experiment conducted on a moderately fertile site; and a whole-tree chamber CO₂ enrichment experiment on an infertile site. The large difference in fertility of the two sites is indicated by a much greater increase in growth at the infertile site (130%) than at the moderate site (15%) when added nutrients augmented the sites to the same optimal fertility⁷. Additional site and soil characteristics that reflect fertility are provided in Methods.

The FACE prototype (FACE_p) was built in 1993 in a 10-year-old, 8.5-m tall loblolly pine plantation in the lower Piedmont plateau, growing on moderately low fertility, acidic clay-loam, at the Duke Forest of Duke University, North Carolina (35° 58' N, 79° 08' W; elevation 130 m). Currently the pines are about 15-m tall, and comprise 98% of the basal area¹⁵. FACE_p has enriched (550 p.p.m.v. CO₂) a 30-m diameter circular patch in the forest since 1994, during daylight hours of the growing season¹⁶. Before CO₂ enrichment in FACE_p commenced, growth (that is, the amount of C sequestered in woody biomass increment) was similar in FACE_p and an adjacent,

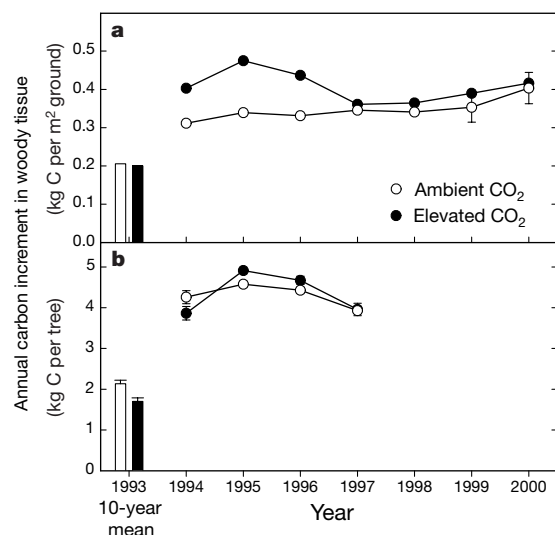


Figure 1 A comparison of annual carbon increment under elevated atmospheric CO₂ concentration (initiated in 1994) and ambient concentration without nutrient addition. **a**, Plot-level comparison between the free-air CO₂ enrichment prototype (FACE_p) and a nearby untreated, ambient CO₂ plot (in the past 2 yr, the number of untreated plots was increased to five). **b**, Individual tree comparison between trees in FACE_p and trees selected at random from the entire stand. Data for 1993 are shown as means for the first 10 yr of the stand's life.

untreated plot (Fig. 1a), as were foliar N concentration and physiological capacity¹⁶.

Averaged over the first 3 yr, the elevated CO₂ plot had a 34% increase in growth relative to the ambient CO₂ plot, reaching a maximum of 40%, but only a 6% increase over the next 4 yr. After four growing seasons of enrichment, we cored all pine trees in FACE_p ($n = 78$) and trees selected at random from the entire stand ($n = 56$; none in the untreated plot), and calculated annual individual tree growth for every year since 1993. The temporal pattern in individual tree growth was fundamentally the same as that of the stand with lower growth of FACE_p trees before the commencement of CO₂ enrichment, a large initial relative response to elevated CO₂, and a subsequent decrease in the response (Fig. 1b).

The CO₂-induced growth enhancement over the first 2 yr in the adjacent replicated FACE study ($n = 3$)¹⁵, which began operation in 1996, averaged about 24%, and ranged from 7 to 57% (1 s.d. = 15%). The initial growth enhancement in the FACE_p plot was not different from that at replicated FACE plots, after matching annual growth for the same year since enrichment commenced in both studies (2 yr, t -test, $P > 0.1$; see Methods for statistical approach). Thus, the early growth responses to elevated CO₂ were consistent in both experiments, but these responses are transient (Fig. 1).

The decrease in response to elevated CO₂ (Fig. 1) might be caused by nutrient limitation that can develop fairly rapidly in this moderate-fertility site with a limited rooting depth (< 0.3 m)¹⁷. The January foliar N concentration was similarly low in FACE_p and in the surrounding stand growing under ambient CO₂ ($0.97 \pm 0.04\%$ by weight; mean $\pm$ s.e., $n = 16$). Plant-soil models predicted responses to an experimental stepwise increase in elevated CO₂ in nutrient-limited forests¹⁸, similar to that shown in Fig. 1. To test whether nutrient limitations reduced the tree response to elevated CO₂, we partitioned FACE_p and its unenriched counterpart in 1998, and added a balanced fertilizer to half the area (> 265 m²)

at the end of the growing season, aiming to attain optimal nutrition (1.4% N)^{7,19}. We also established four plot pairs, randomly positioned in the stand, with one member of each fertilized similarly, thus producing five plot pairs under ambient CO₂. After 1 yr, fertilization increased foliar N concentration under ambient CO₂ to $1.18 \pm 0.08\%$ (mean $\pm$ s.e.), but foliar N remained low under elevated CO₂ (1.02%), similar to unfertilized plots under both ambient and elevated CO₂ ($1.05 \pm 0.07\%$).

Averaged over 1999 and 2000, about 378 ± 16.0 g C m⁻² yr⁻¹ (mean $\pm$ s.e.) was incorporated into woody tissue under ambient CO₂ and without fertilization (Fig. 2a). Annual woody growth under elevated CO₂ without nutrient addition sequestered only 7% more C ($P > 0.05$), and improved nutrient supply in ambient CO₂ atmosphere increased growth by 15% ($P = 0.024$). However, the combination of improved nutrition and elevated CO₂ increased growth by 47% (~ 175 g C m⁻² yr⁻¹; $P = 0.007$)—more than twice the sum of the separate responses to each factor (total ~ 80 g C m⁻² yr⁻¹). This clearly indicates a synergistic effect of CO₂ and nutrient supply that raised the relative CO₂-induced growth response to the maximum previously seen in FACE_p ($\sim 40\%$).

We assessed whether this synergism increases with decreasing site fertility by evaluating the response of trees ($n = 3$, whole-tree open-top chambers) in another North Carolina loblolly pine plantation of a similar age and genetic stock growing on an infertile sandy soil (USDA Forest Service Southeast Tree Research and Education Site, SETRES; $34^{\circ}48'N$, $79^{\circ}12'W$). At this infertile site, foliar N was $1.02 \pm 0.04\%$ (mean $\pm$ s.e.) in both CO₂ environments, increased with fertilization to $1.29 \pm 0.05\%$, and decreased to $1.15 \pm 0.04\%$ when fertilized trees were subjected to elevated CO₂ for two growing seasons (550 p.p.m.v.). Without added nutrients, annual growth did not respond to elevated CO₂ in the 2-yr period ($P > 0.05$; Fig. 2b). Under optimal nutrition¹⁹ and ambient CO₂, the growth of these trees increased by 21% ($P < 0.05$). However, the synergistic enhancement from improved nutrition and higher atmospheric CO₂ was 74% ($P = 0.008$)—more than three times the sum of the separate responses. This suggests that CO₂ responses of growth in pine forests will be highly variable and depend greatly on site fertility, perhaps to the point of not responding at all on the nutritionally poorest sites.

In addition to nutrient limitations, water deficits could also limit forest response to elevated CO₂. The early part of the 1999 growing season was unusually dry along the eastern seaboard of the USA,

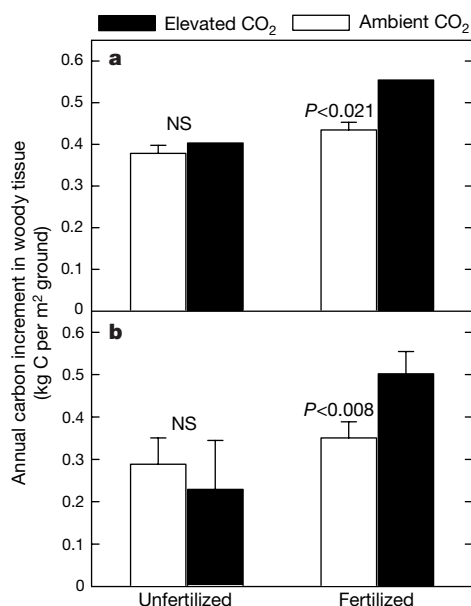


Figure 2 Effect of atmospheric CO₂ concentration and nutrient supply on annual carbon increment in woody tissue of genotypically similar loblolly pine. **a**, A moderate-fertility Duke Forest site (averaged for 1999 and 2000); **b**, an infertile SETRES site (1997 and 1998). P values represent test results between ambient and elevated CO₂ in each fertility level. In both sites, elevating only CO₂ had no significant effect (in **a**, data without nutrient addition are the mean of the past 2 yr shown in Fig. 1a); fertilizing in ambient CO₂ had a significant effect; and fertilizing under elevated CO₂ had significantly higher effects than the sum of the single CO₂ and N effects.

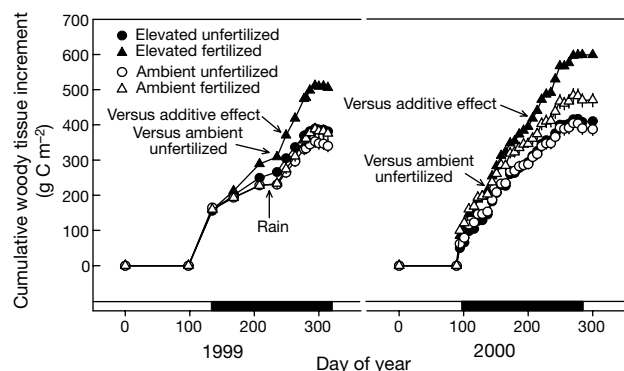


Figure 3 Cumulative carbon increment in woody tissue under ambient and elevated atmospheric CO₂ concentration with and without the addition of nutrients. The carbon increment under elevated CO₂ without nutrient addition was never significantly greater than under ambient CO₂ ($P > 0.05$). Arrows depict the earliest time in each growing season in which the carbon increment under elevated CO₂ with nutrient addition was significantly higher than the indicated treatment ($P < 0.05$). The period of CO₂ enrichment in each year is indicated as a horizontal bar at the bottom. In 1999, drought-breaking rains (indicated) raised soil moisture in the root zone from 0.14 to 0.32 m³ m⁻³. Error bars represent s.e.

with the May–July rainfall 84% below the 30-yr site average. In contrast, rain was ample and evenly distributed during the 2000 growing season. Weekly measurements of stem diameter at the Duke Forest site permitted us to quantify drought-induced interaction effects with the CO₂ and nutrition treatments on the seasonal dynamics of C investment in woody tissue. A fraction of the stem increment that occurred after each rainfall event (such as that in 1999; Fig. 3) was due to actual wood growth and a fraction was due to stem swelling with hydration; we estimated the latter, reversible ‘increment’ as $\leq 17.4 \text{ g C m}^{-2}$ after a given event. Without drought in 2000, growth enhancement in fertilized elevated CO₂ conditions became significant earlier (Fig. 3), and sequestered $45 \text{ g C m}^{-2} \text{ yr}^{-1}$ (27%) more than with early season drought in 1999. Hence, the synergistic effect of CO₂ and nutrients is especially strong when seasonal water limitations are alleviated.

Treatments imposed in a stepwise manner may invoke different responses than gradual changes⁸. Where site fertility is very low, responses to stepwise CO₂ enrichment experiments may be muted (Fig. 2b). Where fertility is moderate, responses may be transient and decrease when soil nutrients are depleted by the initial increase in growth rate¹⁹ (Fig. 1). Thus, the short-term dynamics of the elevated CO₂ response¹⁵ may not portray the longer-term, multi-year responses shown in this study, particularly where dynamic changes in nutrient cycles are involved¹². There are many reasons to expect the long- and short-term responses to be different^{18,20}, and this study highlights the need for caution in broadly extrapolating from short-term results.

Along a gradient of decreasing native-site fertility, nutrient amendments will become increasingly necessary if forest ecosystems (planted or natural) are intended to both take advantage of elevated atmospheric CO₂ and aid in reducing the rate at which atmospheric CO₂ concentrations increase. Under current atmospheric CO₂, nitrogen deposition has been estimated to contribute little to enhanced C sequestration in temperate forests^{10,21,22}, and N deposition may ultimately cause a growth reduction by causing nutritional disharmony²³. Nevertheless, a model comparison has concluded that interaction between CO₂ and N deposition may be involved in the response of terrestrial ecosystems to future levels of atmospheric CO₂ (ref. 24).

Our amendment of $\sim 11 \text{ g N m}^{-2} \text{ yr}^{-1}$ is more than 10 times the current deposition rates of $\sim 0.65 \text{ g N m}^{-2} \text{ yr}^{-1}$ at both sites. At the moderate-fertility site, C sequestration was enhanced by ~ 14.1 and $\sim 17.3 \text{ g C g}^{-1} \text{ N}$ addition under elevated atmospheric CO₂ in the dry and wet year, respectively (that is, an average of $15.7 \text{ g C g}^{-1} \text{ N}$). At the nutritionally poor site, the enhancement under elevated CO₂ was $19.1 \text{ g C g}^{-1} \text{ N}$ addition. This provides a first estimate of the effect of N supply on CO₂-induced C sequestration in woody biomass increment at the scale of forest stands. We note, however, that industrial N fertilizer production and use already has an impact on the global N cycle²⁵, and any plan to enhance C sequestration that relies on a significant increase in N fertilization must be carefully evaluated for both local and global effects. □

Methods

Setting

The FACE and SETRES sites are mid-rotation plantations without density-related mortality. The summers are warm and humid, and winters are moderate. Annual precipitation (1,100–1,200 mm) is evenly distributed throughout the year, but occasional deficits in growing-season water occur. The sites are considered moderate (FACE) and low (SETRES) fertility, as reflected in the heights of dominant *P. taeda* of 15.7 and 10.4 m at age 25 yr, respectively. Soil incubations for net N mineralization at these sites yielded roughly 2.7 (upper 7.5 cm) and 1.3 (15 cm) $\text{g N m}^{-2} \text{ yr}^{-1}$, respectively, and cation exchange capacity was about 12 (upper 15 cm) and 2 (20 cm) cmol (+) g^{-1} . Seasonal CO₂ enrichment in FACE_p, typically from late April to late October, was achieved in 1994–2000 by the free-air CO₂ enrichment technique²⁶.

Annual growth of trees at the FACE_p and its paired untreated plot was measured from 1993 onwards, at the end of each growing season. In 1998, an impermeable barrier to 1-m depth, about three times the depth of fine roots in this forest¹⁷, partitioned both plots into two sections containing roughly the same mass of above-ground biomass per unit ground. Concurrently, four pairs of 10 m × 10 m plots were established nearby with a minimum

distance of 20 m between plots. One randomly selected section (FACE_p and its counterpart) and one randomly chosen plot of each newly established plot pair were fertilized to meet optimal values^{7,19}. An application rate of $11.2 \text{ g N m}^{-2} \text{ yr}^{-1}$ in the form of urea began in July 1998.

The SETRES stand was planted in 1985 in the Sandhills of North Carolina¹⁹, a site with sandy, siliceous, well-drained soil. Annual nutrient additions to meet optimal values commenced in March 1992, with annual applications averaging $11.2 \text{ g N m}^{-2} \text{ yr}^{-1}$ over the 7 yr (ref. 18). Six fertilized trees and six unfertilized trees of a similar diameter range and mean height of $8.6 \pm 1 \text{ m}$ were enclosed in open-top chambers²⁷ in 1996. CO₂ enrichment to $\sim 550 \text{ p.p.m.v.}$ in three chambers per treatment commenced in August 1996, and lasted for two complete growing seasons.

Carbon sequestration

Woody biomass (that is, in branches, bole and roots down to 5 mm in diameter) was estimated at both sites from diameter and allometric equations for ambient CO₂ conditions^{19,28}. At Duke Forest, estimates included trees with diameters $> 50 \text{ mm}$ at 1.4 m above ground, representing 99% of the biomass, less than 5% of which is in hardwood species. At SETRES, equations were available for trees with and without fertilization¹⁹. From SETRES, the density of wood decreased from 0.51 to 0.47 g cm^{-3} with fertilization. From cores taken at FACE_p, the density of wood decreased from 0.52 to 0.48 g cm^{-3} with elevated CO₂. Thus, when growth estimated from diameter was increased by either treatment, the values were adjusted to reflect these reductions. Woody biomass above ground was converted to C content by multiplying by 0.48, and below ground by 0.44, as determined from combustion (CHN Analyzer; Perkin Elmer). In SETRES, biomass was converted to a ground-area value on the basis of 1,260 trees ha^{-1} . At Duke Forest, diameter was measured weekly during the growing season (Fig. 3) with stainless steel dendrometer bands¹⁵ installed 1.4 m above ground on all trees. Edge conditions created when a 3-m wide strip was cleared around FACE_p did not affect growth: there was no pattern in relative growth rate with distance from the edge, nor was there a difference between the population of trees positioned less than one-third of the radius from the periphery and that composed of the rest of the trees ($P > 0.5$).

Statistical considerations

We analysed the SETRES data by paired *t*-test after pairing trees on the basis of initial biomass. Data from FACE_p and its counterpart were not compared statistically (Fig. 1a); however, data from FACE_p and its unenriched counterpart were compared with data from the nearby replicated FACE ($n = 3$), and the 1999 and 2000 data from FACE_p were compared with data from all nearby unenriched plots ($n = 5$) by a *t*-test between a single observation and a population mean²⁹. This approach permits testing whether the response of FACE_p is different from that of the replicated unenriched plots, but without replications it is not possible to estimate the mean response to elevated supplies of both CO₂ and N. Furthermore, we note that the statistical scope of inference of most controlled experiments, including the replicated FACE and SETRES, is limited³⁰.

Received 6 February; accepted 22 March 2001.

1. Tans, P. P. & White, J. W. C. In balance, with a little help from the plants. *Science* **281**, 183–184 (1998).
2. Ciais, P., Tans, P. P., Trolier, M., White, J. W. C. & Francey, R. J. A large Northern Hemisphere terrestrial CO₂ sink indicated by ¹³C/¹²C of atmospheric CO₂. *Science* **269**, 1098–1102 (1995).
3. Schimel, D. S. Terrestrial ecosystems and the carbon cycle. *Global Change Biol.* **1**, 77–91 (1995).
4. Houghton, R. A., Davidson, E. A. & Woodwell, G. M. Missing sinks, feedbacks, and understanding the role of terrestrial ecosystems in the global carbon balance. *Global Biogeochem. Cycles* **12**, 25–34 (1998).
5. Idso, S. B. & Kimball, B. A. Tree growth in carbon-dioxide enriched air and its implications for global carbon cycling and maximum levels of atmospheric CO₂. *Global Biogeochem. Cycles* **7**, 537–555 (1993).
6. Vitousek, P. M. & Howarth, R. W. Nitrogen limitations on land and in the sea: How can it occur? *Biogeochemistry* **13**, 87–115 (1991).
7. Linder, S. Poliar analysis for detecting and correcting nutrient imbalances in Norway spruce. *Ecol. Bull.* **44**, 178–190 (1995).
8. Körner, C. Towards a better experimental basis for upscaling plant responses to elevated CO₂ and climate warming. *Plant Cell Env.* **19**, 1101–1110 (1995).
9. Aber, J. et al. Nitrogen saturation in temperate forest ecosystems. *Bioscience* **48**, 921–934 (1998).
10. Nadelhoffer, K. J. et al. Nitrogen deposition makes a minor contribution to carbon sequestration in temperate forests. *Nature* **398**, 145–148 (1999).
11. Wullschlegel, S. D., Post, W. M. & King, A. W. In *Biotic Feedbacks in the Global Climatic System* (eds Woodwell, G. M. & Mackenzie F. T.) 85–107 (Oxford Univ. Press, New York, 1995).
12. Hu, S., Chapin III, F. S., Firestone, M. K., Field, C. B. & Chiarello, N. R. Nitrogen limitation of microbial decomposition in a grassland under elevated CO₂. *Nature* **409**, 188–191 (2001).
13. Norby, R. J. & Cotrufo, M. F. A question of litter quality. *Nature* **396**, 17–18 (1998).
14. Berntson, G. M. & Bazzaz, F. A. Belowground positive and negative feedbacks on CO₂ growth enhancement. *Plant Soil* **187**, 119–131 (1996).
15. DeLucia, E. H. et al. Net primary production of a forest ecosystem with experimental CO₂ enrichment. *Science* **284**, 1177–1179 (1999).
16. Ellsworth, D., Oren, R., Ce, H., Phillips, N. & Hendrey, G. Leaf and canopy responses to elevated CO₂ in a pine forest using a free-air CO₂ enrichment technique. *Oecologia* **104**, 139–146 (1995).
17. Oren, R., Ewers, B. E., Todd, P., Phillips, N. & Katul, G. Water balance delineates the layer in which soil moisture affects canopy conductance. *Ecol. Appl.* **8**, 990–1002 (1998).
18. Comins, H. N. & McMurtrie, R. E. Long-term responses of nutrient-limited forests to CO₂ enrichment: Equilibrium behavior of plant-soil models. *Ecol. Appl.* **3**, 666–681 (1993).
19. Albaugh, T. J., Allen, H. L., Dougherty, P. M., Kress, L. W. & King, J. S. Leaf area and above- and belowground growth responses of loblolly pine to nutrient and water additions. *For. Sci.* **44**, 1–12 (1998).
20. Bazzaz, F. A., Miao, S. L. & Wayne, P. M. CO₂-induced enhancements of co-occurring tree species decline at different rates. *Oecologia* **96**, 478–482 (1994).
21. Houghton, R. A., Hackler, J. L. & Lawrence, K. T. The U.S. carbon budget: Contributions from land-

- use change. *Science* **285**, 574–578 (1999).
22. Field, C. B. & Fung, I. Y. The not-so-big U.S. carbon sink. *Science* **285**, 544–545 (1999).
23. Oren, R. in *Plant Response to Air Pollution* (eds Iqbal, M. & Yunus, M.) 75–98 (John Wiley & Sons, Chichester, UK, 1996).
24. Kicklighter, D. W. *et al.* A first-order analysis of the potential role of CO₂ fertilization to affect global carbon budget: a comparison of four terrestrial biosphere models. *Tellus B* **51**, 343–366 (1999).
25. Vitousek, P. M. Beyond global warming: Ecology and global change. *Ecology* **75**, 1861–1876 (1994).
26. Hendrey, G. R., Ellsworth, D. S., Lewin, K. F. & Nagy, J. A free-air enrichment system for exposing tall forest vegetation to elevated atmospheric CO₂. *Global Change Biol.* **5**, 293–309 (1999).
27. Tissue, D. T., Thomas, R. B. & Strain, B. R. Growth and photosynthesis of loblolly pine (*Pinus taeda*) after exposure to elevated CO₂ for 19 months in the field. *Tree Physiol.* **16**, 49–60 (1996).
28. Naidu, S. L., DeLucia, E. H. & Thomas, R. B. Contrasting patterns of biomass allocation in dominant and suppressed loblolly pine. *Can. J. For. Res.* **28**, 1116–1124 (1998).
29. Sokal, R. R. & Rohlf, F. J. *Biometry* (W. H. Freeman and Co., San Francisco, 1969).
30. Bolin, B., Canadel, J., Moore, B., Noble, I. & Steffen, W. Effect on the biosphere of elevated atmospheric CO₂. *Science* **285**, 1851–1852 (1999).

Acknowledgements

This study was supported by the Department of Energy through both the Office of Biological and Environmental Research and the National Institute for Global Environmental Change, Southeast Regional Center at the University of Alabama, and by the US Forest Service through both the Southern Global Climate Change Program and the Southern Research Station. This work contributes to the Global Change and Terrestrial Ecosystems (GCTE) core project of the International Geosphere–Biosphere Programme (IGBP). We thank D. E. Pataki for helping with data acquisition; T. Albaugh for descriptive site data; and C. Körner, C. Field and R. Norby for helpful comments on an early version.

Correspondence and requests for materials should be addressed to R.O. (e-mail: ramoren@duke.edu).

An Early Cambrian tunicate from China

D.-G. Shu, L. Chen, J. Han & X.-L. Zhang

Early Life Institute & Department of Geology, Northwest University, Xi'an, 710069, China

Like the Burgess Shales of Canada, the Chengjiang Lagerstätte from the Lower Cambrian of China is renowned for the detailed preservation as fossils of delicate, soft-bodied creatures^{1–9}, providing an insight into the Cambrian explosion. The fossils of possible hemichordate chordates^{5–7} and vertebrates⁹ have attracted particular attention. Tunicates, or urochordates, comprise the most basal chordate clade¹⁰, and details of their evolution could be important in understanding the sequence of character acquisition that led to the emergence of chordates and vertebrates^{11–18}. However, definitive fossils of tunicates from the Cambrian are scarce or debatable^{4,9,19–24}. Here we report a probable tunicate *Cheungkongella ancestralis* from the Chengjiang fauna. It resembles the extant ascidian tunicate genus *Styela* whose morphology could be useful in understanding the origin of the vertebrates.

Phylum Chordata
Subphylum Urochordata
Class Ascidiacea
Cheungkongella ancestralis gen. et sp. nov.

Type species. *Cheungkongella ancestralis*.

Etymology. Genetic name is a metaphor of China and is also in honour of the Cheungkong Scholars Programme that supports this work; the specific name is a reference to its possible primitive position.

Holotype. Early Life Institute (ELI), Northwest University, Xi'an: ELI-0000195.

Stratigraphy and locality. Qiongzhusi Formation, Yu'an-shan Member (*Eoredlichia* Zone); Lower Cambrian. The specimen was collected by L.C. and J.H. from the same locality and horizon as the

animal *Xidazoon*⁸ and agnathan vertebrate *Myllokunmingia*⁹.

Diagnosis. The body is club-shaped, reminiscent of extant ascidian *Styela*, with two-fold division: an upper main body and a lower thick supporting stem attached to hard substratum (Fig. 1). The body is wholly enclosed within a structure interpreted as a secreted tunic. The stem tapers downward, and the main body is bucket-shaped in outline, bearing a large oral siphon with short tentacles on its top and a small cloacal one on the lower dorsal side. A pharynx occupies over two-thirds of the body volume.

Description. *Cheungkongella ancestralis*, new genus and species, is known from a single specimen, with a total length of about 25 mm. The whole body consists of two regions: a stout stem, which in life supported a sub-spherical main body. The stem (about 15 mm long) tapers distally, and is attached to the exterior surface of the left, free cheek of a trilobite *Eoredlichia intermedia*, an index fossil for the Lower Cambrian. The stem bears some transverse creases, consistent with an enclosing tunic, and prominent longitudinal 'ribs'. The distal section has a conspicuously coarse texture, and has several patches of agglutinated sediment including quartz grains.

The main body (roughly 10 mm long) was probably sub-spherical in life. Wrinkling of the compressed body on the ventral side (opposite to cloacal siphon) is consistent with folding of a tough

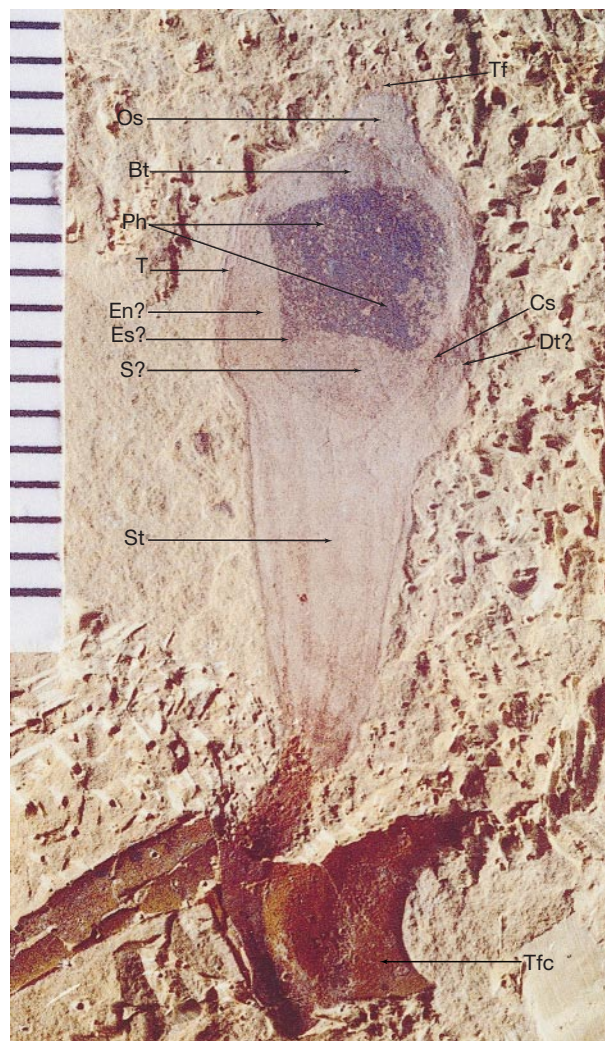


Figure 1 The Lower Cambrian urochordate *Cheungkongella ancestralis* gen. et sp. nov. from Haikou, Kunming, Yunnan. Specimen ELI-0000195, viewed from the left. Scale bar, 1 mm. Bt, buccal tentacles; Cs, cloacal siphon; Dt?, degenerating tail; Os, oral siphon; En?, presumed endostyle; Es?, possible esophagus; Ph, pharynx; S?, presumed stomach; St, stem; T, tunic; Tf, tentacle-like fringe; Tfc, trilobite free cheek.

tunic. The large funnel-like structure filled with a layer of sediment on the top is interpreted as the oral siphon, as is found in all solitary ascidians. To the lower right of the main body, presumably representing the original dorsal side, a narrow but distinct extension is interpreted as the excurrent (or cloacal) siphon, and is consistent with a rather marked change in level in this area. Lower and outside to the cloacal siphon there is an arcuate area, where there is a recognizable recurvate dark ridge. This structure is difficult to interpret, but when compared to the last stage of metamorphosis of living ascidians²⁵, it could represent the remnant of a larval tail.

A prominent rectangular dark area, covered with numerous lighter dots, is located in the upper part of the body cavity. The lower ventral corner of the dark area extends downward as a 'tail-like' structure. The dark area, compared to modern ascidians, with its dextral configuration, large size and appropriate location in the body, is consistent with its identification as the pharynx. Whether the lighter dots represent gill openings remains to be confirmed.

Two interesting structures are located above and below the oral siphon. The area between the siphon and the pharynx is preserved in dark grey. Observation shows a longitudinal alignment of structures suggestive of buccal tentacles. Above the oral siphon, another set of short tentacle-like filaments is recognizable. They are superficially similar to siphonal fringe or oral lobes of some extant ascidians, but are also reminiscent of the tentacles of extant lophophorate phoronids and the Lower Cambrian lophophorate *Cambrotentacus*⁴. We suggest that this Cambrian tunicate was a suspension feeder, with water entering the oral siphon and being expelled through the cloacal siphon after filtration.

The main body and upper two-thirds of the supporting stem lie laterally on the same bedding plane, but the lower third of the stem is bent steeply into the sediment and attached to a free cheek of a trilobite. This arrangement could indicate burial of the tunicate *in situ*. The presence of agglutinated quartz grains, substantially coarser than the surrounding matrix, on the lower stem suggests, however, that the animal inhabited a higher-energy, sandy sea floor and was transported to its point of burial. During deposition the heavier trilobite sclerite sank first, so tethering the tunicate in the rapidly accumulating sediment. The three-dimensional preservation and remains of the delicate tentacles are indicative of its suffering little decomposition.

Urochordates are believed to represent the most basal chordate branch within Chordata^{11,26}; however, whether the ancestral chordates were free-swimming or sessile has been a long-standing question^{18,26,27}. Traditional hypotheses hold that vertebrates evolved by pedomorphosis from a urochordate-like larval stage, and that the ancestor of chordates would have resembled a sessile lophophorate^{12,13,27}. Recent models, supported by molecular data, posit a free-swimming ancestry of chordates, including urochordates^{28–30}.

Fossils may preserve combinations of characters not seen in extant groups, and so are crucial for testing schemes of how characters were acquired in the origin of new body plans. The interpretation of the present specimen, as possessing oral tentacles comparable to those seen in lophophorates, is consistent with traditional views—if not modern, molecule-based hypotheses—but a single example is far from being conclusive. Further palaeontological and molecular work is needed to investigate the problem. □

Received 21 November 2000; accepted 15 March 2001.

- Shu, D.-G., Geyer, G., Chen, L. & Zhang, X.-L. Redlichian trilobites with preserved soft-parts from the Lower Cambrian Chengjiang fauna. *Beringia Spec. Iss.* **2**, 203–241 (1995).
- Shu, D.-G., Zhang, X. & Geyer, G. Anatomy and systematic affinities of Lower Cambrian bivalved arthropod *Isoxys auritus*. *Alcheringa* **19**, 333–342 (1995).
- Hou, X. & Bergstrom, J. Arthropods of the Lower Cambrian Chengjiang fauna, southwest China. *Fossils Strata* **45**, 1–115 (1997).
- Zhang, X.-L., Shu, D.-G., Li, Y. & Han, J. New sites of Chengjiang fossils: crucial windows on the Cambrian explosion. *J. Geol. Soc. Lond.* **158**, 211–218 (2001).
- Chen, J.-Y., Dzik, J., Edgecombe, G. D., Ramskoeld, L. & Zhou, G.-Q. A possible early Cambrian chordate. *Nature* **377**, 720–722 (1995).

- Shu, D.-G. & Zhang, X.-L. Reinterpretation of *Yunnanozoon* as the earliest known hemichordate. *Nature* **380**, 428–430 (1996).
- Shu, D.-G., Conway Morris, S. & Zhang, X.-L. A *Pikaia*-like chordate from the Lower Cambrian of China. *Nature* **384**, 156–157 (1996).
- Shu, D. *et al.* A pipiscid-like fossil from the Lower Cambrian of South China. *Nature* **400**, 746–749 (1999).
- Shu, D. *et al.* Lower Cambrian vertebrates from South China. *Nature* **402**, 42–46 (1999).
- Cameron, C. B., Garey, J. R. & Swalla, B. J. Evolution of the chordate body plan: new insights from phylogenetic analyses of deuterostome phyla. *Proc. Natl Acad. Sci. USA* **97**, 4469–4474 (2000).
- Garstang, W. The morphology of the Tunicata and its bearing on the phylogeny of the Chordata. *J. Microscop. Soc.* **72**, 51–87 (1928).
- Romer, A. S. *The Vertebrate Story* (Univ. Chicago Press, 1971).
- Gee, H. *Before the Backbone: Views on the Origins of Vertebrates* (Chapman and Hall, London, 1996).
- Ogasawara, M., Di Lauro, R. & Satoh, N. Ascidian homologs of mammalian thyroid transcription Factor-1 gene expressed in the endostyle. *Zool. Sci.* **16**, 559–565 (1999).
- De Gregorio, A. & Levine, M. Ascidian embryogenesis and the origins of the chordate body plan. *Curr. Opin. Genet. Dev.* **8**, 457–463 (1998).
- Swalla, B. J. *et al.* Urochordates are monophyletic within the deuterostomes. *System. Biol.* **49**, 52–64 (2000).
- Nielsen, C. *Animal Evolution: Interrelationships of Living Phyla* (Oxford Univ. Press, Oxford, 1997).
- Conway Morris, S. in *Atlas of the Burgess Shale* (ed. Conway Morris, S.) **26** (Palaeontological Association, London, 1982).
- Briggs, D. E. G. *et al.* The Fossils of the Burgess Shale (Smithsonian, Washington, 1994).
- Satoh, N. *Developmental Biology of Ascidians* (Cambridge Univ. Press, New York, 1994).
- Mueller, K. J. *Palaebotryllus* from the Upper Cambrian of Nevada—a possible ascidian. *Lethaia* **10**, 107–118 (1977).
- Lehnert, O., Miller, J. F. & Cochrane, K. Alaeobotryllus and friends: Cambro-Ordovician record of probable ascidian tunicates. *Acta Univ. Carol. Geol.* **43**, 447–450 (1999).
- Lohmann, H. Oesia disjuncta Walcott, eine Appendicularie aus dem Kambrium. *Mitt. Zool. Mus. Hamburg* **38**, 69–75 (1922).
- Zhang, A. Fossil appendicularians in the early Cambrian. *Scient. Sinica B* **30**, 888–896 (1987).
- Meglitsch, P. A. & Schram, F. P. *Invertebrate Zoology* 3rd edn, 576–587 (Oxford Univ. Press, 1991).
- Wada, H. & Satoh, N. Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl Acad. Sci. USA* **91**, 1801–1804 (1994).
- Berrill, N. J. *The Origin of Vertebrates* (Oxford Univ. Press, Oxford, 1955).
- Carter, G. S. Chordate Phylogeny. *Syst. Zool.* **6**, 187–192 (1957).
- Jollie, M. What are the 'Calcichordata'? and the larger question of the origin of chordates. *Zool. J. Linn. Soc.* **75**, 167–188 (1982).
- Wada, H. Evolutionary history of free-swimming and sessile lifestyles in urochordates as deduced from 18S rDNA molecular phylogeny. *Mol. Biol. Evol.* **15**, 1189–1194 (1998).

Acknowledgements

This work is supported by the Ministry of Sciences and Technology of China, National Natural Science Foundation of China, Ministry of Education of China and Cheung Kong Scholars Programme, Sciences and Technology Committee of Shaanxi Province, and National Geographic Society, USA. We are grateful to B. J. Swalla and S. Conway Morris for suggestions. We thank N. Satoh, K. Yasui, H. Wada, R. P. S. Jefferies and S. M. Shimeld for remarks. We also thank Z. Luo, Z. Zhang, L. Guo, F. Yao and H. Liu for technical assistance.

Correspondence and requests for materials should be addressed to D.-G.S. (e-mail: dgshu@sein.sxgb.com.cn).

Orientation-selective adaptation and tilt after-effect from invisible patterns

Sheng He* & Donald I. A. MacLeod†

* Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455, USA

† Department of Psychology, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

Exposure to visual patterns of high contrast (for example, gratings formed by alternating white and black bars) creates after-effects in perception. We become temporarily insensitive to faint test patterns that resemble the pre-exposed pattern (such as gratings of the same orientation), and we require more contrast to detect them¹. Moreover, if the test pattern is slightly tilted relative to the pre-exposed one, this tilt may be perceptually exaggerated: we experience a tilt after-effect^{2,3}. Here we show that these visual after-effects occur even if the pre-exposed grating is too fine to be perceptually resolved. After looking at a very fine grating, so high

in spatial frequency that it was perceptually indistinguishable from a uniform field, observers required more contrast to detect a test grating presented at the same orientation than one presented at the orthogonal orientation. They also experienced a tilt after-effect that depended on the relation of the test pattern's tilt to the unseen orientation of the pre-exposed pattern. Because these after-effects are due to changes in orientation-sensitive mechanisms in visual cortex⁴⁻⁶, our observations imply that extremely fine details, even those too fine to be seen, can penetrate the visual system as far as the cortex, where they are represented neurally without conscious awareness.

Vision is not perfect. When grating patterns are used as resolution targets, the highest spatial frequency resolvable by humans is about 50 cycles per degree of angle subtended at the observer's eye⁷. The higher spatial frequencies are lost partly because of diffraction and other sources of optical blur. These optical losses can be bypassed if the targets are presented as interference fringe patterns generated directly on the retina, but when this is done the resolution limit is only slightly increased⁷⁻¹⁰. It follows that there are neural, as well as optical, limitations on visual resolving power; but the roles of the retina and the brain in imposing these limitations have not been worked out. Here we report observations suggesting that perceptually unresolved patterns can penetrate to the visual cortex, and selectively modify visual sensitivity and pattern perception there. Our failure to perceive these patterns consciously is therefore partly

caused by limitations in cortical processing.

Orientation information is first analysed at the primary visual cortex. Neurons in the primary visual cortex respond selectively to their preferred orientation⁶. Moreover, prolonged exposure of such neurons to a grating in one orientation will selectively reduce their population sensitivity to gratings in the same orientation^{4,5,11}. This property of the cortical neurons made it possible for us to test psychophysically whether a perceptually unresolvable grating could still be represented in visual cortex. The logic is simple: if an unresolvable grating can produce an orientation-selective perceptual after-effect, then its spatial structure, including its orientation, is registered at least at the earliest stages of cortical processing.

Using a He-Ne laser interferometer, subjects in our first experiment viewed a horizontal or vertical adapting grating at full contrast for one minute. Next, their ability to detect horizontal or vertical test gratings of various contrasts was assessed (Fig. 1a). Every 6 s, the continued presentation of the adapting grating was interrupted by a 1-s trial in which either a horizontal or a vertical test grating was presented, the orientation being randomly varied from trial to trial. During each trial, the test grating was presented during either the first or (randomly) the second of two successive 250-ms intervals demarcated by audible clicks. After each trial, the observer had to indicate whether it was the first or the second interval that had contained the test grating. The 500-ms pair of test intervals were preceded and followed by a 250-ms period during which the field

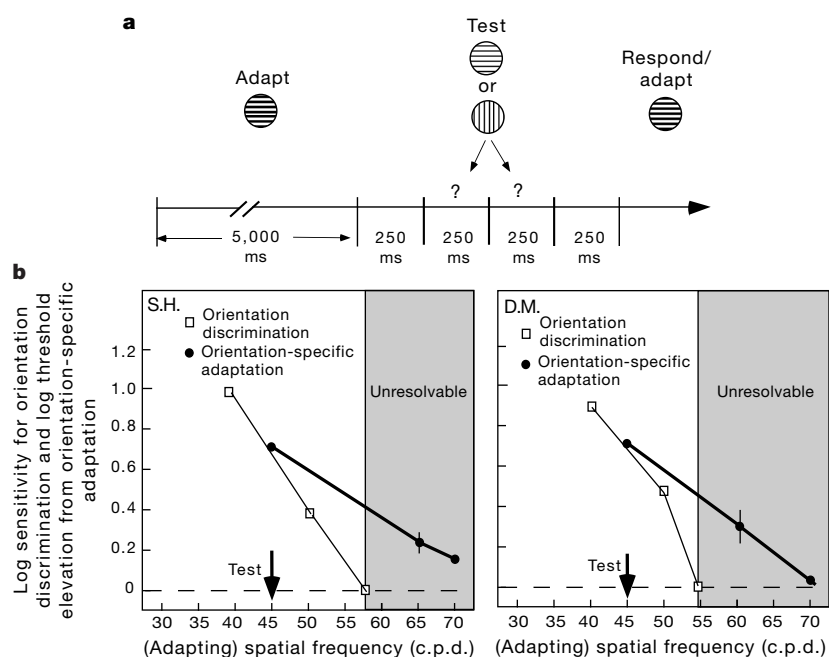


Figure 1 Orientation-selective adaptation experiment. **a**, Experimental procedure. After viewing a horizontal or vertical adapting grating at full contrast for 1 min, observers' sensitivity to briefly presented testing gratings were measured in a 2AFC staircase procedure. The procedure searched for threshold contrasts for a 45-c.p.d. horizontal and vertical testing grating in two independent but interleaved staircases. Both adapting and testing gratings were 3° in diameter with a mean luminance of 1,000 trolands. Shown is the sequence of one 6-s trial, during which the test grating was presented during either the first or (randomly) the second of two successive 250-ms intervals demarcated by audible clicks. The test grating could be either horizontal or vertical—the orientation was varied randomly from trial to trial. After each trial, the observer had to indicate whether it was the first or the second interval that had contained the test grating. The 500-ms pair of test intervals were preceded and followed by a 250-ms period in which the field was uniform; during the remaining 5 s of the cycle time, the field was again occupied by the adapting grating. **b**, Comparison of the effect of orientation-specific adaptation with the sensitivity of orientation discrimination, both plotted as functions of spatial frequency. Open squares plot the sensitivities of the two observers discriminating between two

grating orientations. For both observers, this sensitivity drops rapidly as the spatial frequency approaches 50 c.p.d. S.H. could not discern the orientation of gratings beyond 57 c.p.d.; D.M. could not discern that beyond 55 c.p.d. In comparison, the orientation-specific adaptation effect decreases more slowly as the spatial frequency increases, still measurable beyond the point at which the grating is no longer resolved. For both subjects, adaptation to a grating had different effects on test gratings of horizontal and vertical orientations, depending on the orientation of the adapting grating. The threshold contrast for the test gratings of the same orientation as the adapting grating was elevated relative to that for perpendicular test gratings. The magnitude of the orientation-selective adaptation after-effect was expressed as the difference in relative sensitivity between the two test orientations when horizontal and vertical adaptation were compared. Filled circles are the logarithmic difference between the relative sensitivities for horizontal and for vertical pre-adaptation. Thresholds were elevated for the same adapt/test orientations, and reduced for the orthogonal adapt/test orientation. This is true even when the adapting frequency far exceeds the conventional resolution limit (shaded areas). Error bars, $\pm$ 1 s.e.

was uniform; during the remaining 5 s of the cycle time, the field was again occupied by the adapting grating. Space average intensity remained constant throughout at 1,000 trolands.

The test grating was always 45 cycles per degree (c.p.d.), a spatial frequency just below the resolution limit. The spatial frequency of the adapting grating was varied in different runs from the visible to the invisible range; at adapting frequencies higher than about 55 c.p.d., the subjects were unable to tell the orientation of the adapting grating in a forced-choice discrimination task (Fig. 1b).

For both subjects tested (S.H. and D.M.), pre-adaptation to a grating had different effects on test gratings of horizontal and vertical orientations. The threshold contrast for the test gratings of the same orientation as the adapting grating was elevated relative to that for perpendicular test gratings. Figure 1b shows the magnitude of the orientation-selective adaptation after-effect, expressed as the difference in relative sensitivity between the two test orientations when horizontal and vertical adaptation are compared. The figure plots the logarithmic difference between the relative sensitivities for horizontal and for vertical pre-adaptation.

Orientation-selective adaptation using resolvable gratings has been shown long ago¹, but it is surprising that gratings at and even beyond the resolution limit can still produce orientation-selective adaptation effects: conscious perception is unresponsive or poorly responsive to such spatial frequencies, with visual sensitivity declining as the fifth to sixth power of spatial frequency in this range¹⁰. But our 45-c.p.d. interference fringe patterns produced effects comparable in magnitude to those documented for easily visible spatial frequencies; and even invisibly fine adapting frequencies (60–70 c.p.d.) produced measurable effects. Unexpectedly, in comparison with pre-adaptation to uniform fields, pre-adaptation to unresolvably fine patterns actually improved contrast sensitivity for test gratings very near the limit of resolution; this effect, not yet well understood, was not orientation selective and is factored out in Fig. 1b. A more detailed description of the results can be found in Supplementary Information.

These results imply that gratings too fine to be resolved can nevertheless selectively activate orientation-selective functional units (neurons or synapses) in the cortex. Orientation information must be registered at least at the cortical stage of processing where orientation-specific adaptation arises; and it must then be lost to perception owing to cortical limitations on spatial resolution.

As a further test of this idea, we asked whether a tilt after-effect is experienced after adaptation to unresolvable gratings. It is generally thought that perception of local orientation is based on the distribution of activity among many orientation-selective neurons in visual cortex, and that adaptation at a particular orientation selectively depresses the sensitivity of neurons sensitive to that orientation. When these neurons are subsequently stimulated with a slightly different orientation than the adapting one, the activity distribution is shifted away from the adapting orientation; as a result, the test grating appears to be tilted further away from the adapting orientation—the tilt after-effect¹². In this view, the tilt after-effect implies that the functional units whose sensitivity is modulated by adaptation are not only orientation selective but are actually used for representing orientation in perceptual experience¹³.

To determine whether unresolvable gratings can activate these orientation-selective and orientation-signalling cortical units, subjects were asked to set a briefly presented 48 c.p.d. grating to a subjectively horizontal orientation after adapting to a full-contrast grating with a spatial frequency at least 12% higher than the resolution limit (66 c.p.d. for S.H., and 60 c.p.d. for D.M.). The adapting grating was tilted away from horizontal by 15° clockwise or anti-clockwise (see Fig. 2a for the procedure). With visible gratings, the perceived orientation of a test grating close in orientation to the adapting orientation is pushed away from the adapting orientation.

If an unresolvable grating similarly induces a tilt after-effect, for the test grating to be perceived as horizontal, it will need to be slightly tilted clockwise after adapting to a 15° clockwise-tilted grating, and slightly anti-clockwise after adapting to an anti-clockwise grating. This is exactly what we observed. Figure 2b shows the mean tilt after-effect (offset from horizontal) after adapting to an unresolvable grating. The magnitude of the tilt after-effect is comparable to what has been observed with a high-contrast resolvable grating³. This effect was maintained at an adapting frequency of 66 c.p.d., which yielded a tilt after-effect of $1.2 \pm 0.28^\circ$ (mean $\pm$ s.e.m.) in each direction. In these experiments, the two pre-adapting tilts were presented in randomly interleaved blocks in each run, so that the subject had no knowledge of the adapting tilt when the adapting stimulus was unresolvable.

Our results indicate that gratings of high spatial frequency may be represented in the cortex (and generate after-effects) even when not consciously perceived¹⁴. The cortex could limit perceptual resolution either by blocking high-spatial-frequency signals after the stage at which the after-effects are generated, or by blocking weak signals of any spatial frequency. To decide between these possibilities, we tried a low-spatial-frequency adapting grating equated in contrast with the unresolvable one relative to their respective thresholds. Near the resolution limit, the threshold contrast was found to vary at least as the fifth power of frequency¹⁰. On that basis, the unresolvable adapting gratings had roughly half the (extrapolated) threshold contrast. For comparison, we chose an adapting grating of 36 c.p.d., placed roughly symmetrically with the unresolvable one

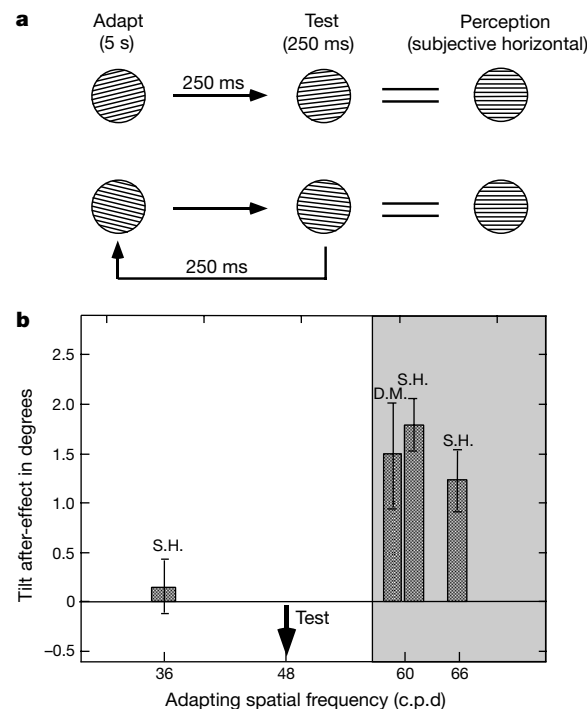


Figure 2 Tilt after-effect experiment. **a**, Experimental procedure. The observers' task was to set the test grating to the perceptual horizontal after adapting to a grating that was tilted either clockwise or anti-clockwise 15° from horizontal. With a resolvable adapting grating, observers set the test grating slightly clockwise when the adapting grating was clockwise, and anti-clockwise when the adapting was anti-clockwise. Here, unresolvable gratings also produced a tilt after-effect. See also Supplementary Information. **b**, Settings for perceptual horizontal after adapting to slightly tilted unresolvable gratings (in the shaded frequency range), or to a subthreshold resolvable grating equated for detectability as a control. Observers' settings for horizontal were consistently tilted towards the adapting orientation, even though the adapting grating was invisible, and observers had no idea what its orientation was. For the control condition at 36 c.p.d., observers set the horizontal veridically. Error bars, ± 1 s.e.

about the test grating (48 c.p.d.) on the frequency scale. The contrast of the 36 c.p.d. adapting grating was set to half of its threshold contrast, 3.5%. No measurable tilt after-effect was found under this condition (Fig. 2b).

Similarly, in measurements of the elevation in contrast threshold produced by pre-exposure to a range of adapting grating contrasts, at three spatial frequencies spanning the resolution limit, we found that adapting gratings of frequency greater than the resolution limit were more effective than correspondingly subthreshold gratings of lower frequency (although subthreshold gratings in the high but resolvable frequency range were not entirely ineffective). Thus, the cortical requirement for conscious perception seems to depend on spatial frequency, and not only on contrast or signal strength represented at the cortical input.

The idea that limits on visual resolution are partly imposed at the cortical level is supported by evidence that cortically projecting thalamic relay neurons in macaque often respond well to spatial frequencies far above the human resolution limit, in some cases exceeding 100 c.p.d. (ref. 15). If the projection from thalamus to cortex were as precise as the one from retina to thalamus, this would allow the visual system to form a representation of unresolvable patterns at the cortical site of pattern adaptation. The lower spatial-frequency limits for cortical after-effects (70 c.p.d.), as compared with thalamic neurons (100 c.p.d.), may reflect imperfect precision in the projection from thalamus to cortex¹⁶.

In normal vision with incoherent light, diffraction markedly reduces the retinal image contrast for spatial frequencies near the resolution limit. Why should the cortex have orientation-selective mechanisms (or frequency-selective ones¹⁷) that respond to high spatial frequencies that are normally only faintly represented in the retinal image? One answer derives from the view¹⁸ that neural mechanisms might compensate for optical blur: frequency components that have been optically attenuated might even require a reciprocal enhancement of neural sensitivity for their appropriate representation.

But on our evidence, activation of orientation-selective units at the stage of cortical pattern adaptation is not sufficient for perceptual awareness of the pattern orientation. The nature of the added requirement is not clear: one possibility¹⁹ is that information must be relayed from primary visual cortex to another region of the brain to be represented in conscious experience. □

Received 24 January; accepted 28 March 2001.

1. Blakemore, C. & Campbell, F. W. Adaptation to spatial stimuli. *J. Physiol. (Lond.)* **200**, 11P–13P (1969).
2. Gibson, J. J. & Radner, M. Adaptation, aftereffect and contrast in the perception of tilted lines. *J. Exp. Psychol.* **20**, 453–467 (1937).
3. Wolfe, J. M. Short test flashes produce large tilt aftereffects. *Vis. Res.* **24**, 1959–1964 (1984).
4. Maffei, L., Fiorentini, A. & Bisti, S. Neural correlate of perceptual adaptation to gratings. *Science* **182**, 1036–1038 (1973).
5. Movshon, J. A. & Lennie, P. Pattern-selective adaptation in visual cortical neurones. *Nature* **278**, 850–852 (1979).
6. Hubel, D. H. & Wiesel, T. N. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B* **198**, 1–59 (1977).
7. Campbell, F. W. & Gubisch, R. W. Optical quality of the human eye. *J. Physiol. (Lond.)* **186**, 558–578 (1966).
8. Le Grand, Y. La formation des images rétiniennes. *Rev. Opt.* **16**, 201–241 (1937).
9. Sekiguchi, N., Williams, D. R. & Brainard, D. H. Efficiency in detection of isoluminant and isochromatic interference fringes. *J. Opt. Soc. Am. A* **10**, 2118–2133 (1993).
10. He, S. & MacLeod, D. I. A. Local luminance nonlinearity and receptor aliasing in the detection of high-frequency gratings. *J. Opt. Soc. Am. A* **13**, 1139–1151 (1996).
11. Carandini, M., Movshon, J. A. & Ferster, D. Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology* **37**, 501–511 (1998).
12. Coltheart, M. Visual feature-analysers and aftereffects of tilt and curvature. *Psychol. Rev.* **78**, 114–121 (1971).
13. Barlow, H. B. Single units and sensation: A neuron doctrine for perceptual psychology? *Perception* **1**, 371–394 (1972).
14. Levi, D. M., Harwerth, R. S. & Smith, E. L. Humans deprived of normal binocular vision have binocular interactions tuned to size and orientation. *Science* **206**, 852–854 (1979).
15. McMahon, M. J., Lankheet, M. J. M., Lennie, P. & Williams, D. R. Fine structure of parvocellular receptive fields in the primate fovea revealed by laser interferometry. *J. Neurosci.* **20**, 2043–2053 (2000).
16. Reid, R. C. & Alonso, J. M. Specificity of monosynaptic connections from thalamus to visual cortex. *Nature* **378**, 281–284 (1995).

17. Smallman, H. S., MacLeod, D. I. A., He, S. & Kenridge, R. W. Fine grain of the neural representation of human spatial vision. *J. Neurosci.* **16**, 1852–1859 (1996).
18. Ratliff, F. *Mach Bands: Quantitative Studies on Neural Networks in the Retina* 1–365 (Holden-Day Inc., London, 1965).
19. Crick, F. & Koch, C. Are we aware of neural activity in primary visual cortex? *Nature* **375**, 121–123 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank H. Smallman for his help during the experiment. This work is supported by grants from NIH and Alfred Sloan Foundation.

Correspondence and requests for materials should be addressed to S.H. (e-mail: sheng@tc.umn.edu).

Disruption of neurotransmission in *Drosophila* mushroom body blocks retrieval but not acquisition of memory

Josh Dubnau*, Lori Grady*, Toshi Kitamoto† & Tim Tully*

* Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA

† Division of Neurosciences, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, California 91010, USA

Surgical, pharmacological and genetic lesion studies have revealed distinct anatomical sites involved with different forms of learning. Studies of patients with localized brain damage and work in rodent model systems, for example, have shown that the hippocampal formation participates in acquisition of declarative tasks but is not the site of their long-term storage^{1,2}. Such lesions are usually irreversible, however, which has limited their use for dissecting the temporal processes of acquisition, storage and retrieval of memories^{3,4}. Studies in bees and flies have similarly revealed a distinct anatomical region of the insect brain, the mushroom body, that is involved specifically in olfactory associative learning^{5,6}. We have used a temperature-sensitive *dynamitin* transgene, which disrupts synaptic transmission reversibly and on the time-scale of minutes⁷, to investigate the temporal requirements for ongoing neural activity during memory formation. Here we show that synaptic transmission from mushroom body neurons is required during memory retrieval but not during acquisition or storage. We propose that the hebbian processes underlying olfactory associative learning reside in mushroom body dendrites or upstream of the mushroom body and that the resulting alterations in synaptic strength modulate mushroom body output during memory retrieval.

The mushroom body is a central control neuropil that receives multimodal input^{8–10}. In *Drosophila*, one hemisegment of the mushroom body consists of around 2,500 kenyon cells, whose primary afferents convey olfactory input through the antennal-glomerular tract (AGT; Fig. 1). The AGT projects from the antennal lobe, which itself receives olfactory input from the antennae. Mushroom body efferents project to other neuropil regions that are ultimately involved in motor output. Mushroom body neurons are believed to integrate multimodal information, including olfactory stimuli, and to modulate behavioural responses through motor output. Consistent with this anatomical view of the mushroom body, genetic and pharmacological disruptions of mushroom body

neurons interfere with olfactory associative learning but do not perturb sensorimotor responses to the odours or footshock used in the conditioning assay^{5,11}. Normal cyclic AMP signalling in mushroom body neurons is required for olfactory associative learning^{11,12}, indicating that molecular mechanisms in mushroom body neurons may be involved in behavioural plasticity.

Although these findings show that the mushroom body participates in olfactory learning, they do not distinguish among the temporal stages of memory formation (acquisition, storage and retrieval). We have used a mutant transgene of *dynammin* to gain temporal and spatial control of neuronal activity *in vivo*⁷. Mutations in *dynammin* were first discovered in a screen for temperature-sensitive paralytic mutants in *Drosophila*¹³. *shibire* (*shi*) mutants were identified that displayed normal locomotor activity at permissive temperature (20 °C) but became paralysed soon after a shift to restrictive temperature (30 °C). These *shi*^{ts} mutants recovered within minutes after they were shifted back to permissive temperature. Analysis of ultrastructure and function at the neuromuscular junction of *shi*^{ts} mutants has revealed a reversible, temperature-dependent loss of synaptic transmission concomitant with depletion of synaptic vesicles^{14–16}. The *shi* transcription unit encodes the fly homologue of dynamin^{17,18}, a synaptic GTPase required for fission of endocytic vesicles from the plasma membrane. Molecular genetic analysis has indicated that dynamin may assemble into multimeric complexes and that missense mutations in the GTPase domain can affect dynamin function in a dominant-negative fashion^{19,20}. This finding is consistent with the observation that several mutant alleles affecting this protein domain (for example, *shi*^{ts1}) are genetically semidominant.

We expressed a *shi*^{ts1} mutant complementary DNA⁷ in transgenic animals. When overexpressed in a wild-type background, this *UAS-shi*^{ts1} transgene reversibly interferes with neuronal transmission in a temperature-dependent, dominant-negative fashion. We used the Gal4 enhancer-trap system²¹ to manipulate the spatial expression of *UAS-shi*^{ts1}. Analysis of multiple Gal4 drivers has established a panel of lines with preferential expression in mushroom body neurons. These Gal4 driver lines have been used extensively to discern the neural architecture of the mushroom body, direct transgenic expression to the mushroom body for genetic rescue experiments and assess the role of cAMP signalling in the mushroom body during olfactory learning^{11,12,22}. We have built upon this genetic approach to dissociate the spatial and temporal requirements for

synaptic transmission in the mushroom body during acquisition, storage and retrieval of olfactory memory.

Widespread expression of the *UAS-shi*^{ts1} transgene reproduces the reversible temperature-dependent paralysis of *shi*^{ts1} mutants. The MJ85b enhancer line drives Gal4 expression in all neurons of the adult brain except photoreceptors (Fig. 2a)²³. When *UAS-shi*^{ts1} is expressed with this Gal4 driver, transgenic flies rapidly become paralysed after being shifted to restrictive temperature and then recover their locomotion within one minute after a shift back to permissive temperature (Fig. 2c and data not shown)²⁴.

When expression of the *UAS-shi*^{ts1} transgene was limited predominantly to the mushroom body (Fig. 2b), locomotion was normal (Fig. 2d), as were sensorimotor responses to the odours and footshock stimuli used in pavlovian learning assays (Table 1). For these experiments we used two well characterized enhancer-trap lines, C747 and C309. These GAL4 enhancers have been used to drive expression of a dominant-negative form of G_{α} in mushroom body neurons. This disruption abolishes olfactory learning and also has no effect on sensorimotor responses to the odours or footshock¹¹.

We next studied the associative learning produced by temporally pairing odour and footshock in a pavlovian conditioning

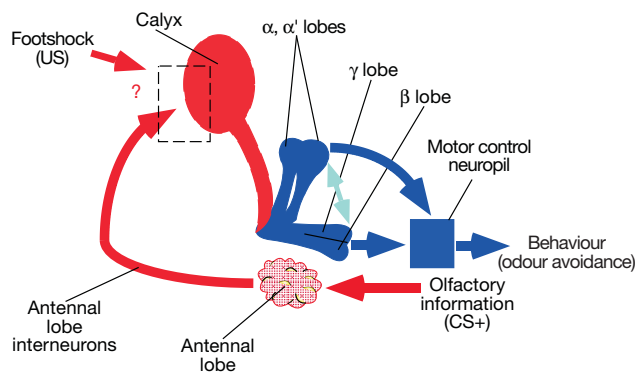


Figure 1 Neural model of olfactory learning in *Drosophila*. The mushroom body is a central control neuropil that receives multimodal input (red) and sends output (blue) to other brain regions involved in the motor control of odour avoidance responses. Feedback within mushroom body neurons (light blue arrow) is likely. Olfactory information is relayed from antennae to mushroom body dendrites (calyx) by antennal lobe interneurons. Concomitant input from footshock through an unknown circuit (?) presumably induces a hebbian process in the mushroom body calyx, thereby modulating synaptic transmission from mushroom body axons (the α -, α' -, β - and γ -lobes) to motor circuits that produce odour avoidance responses. We propose that the hebbian processes underlying olfactory associative learning occur in mushroom body dendrites (box).

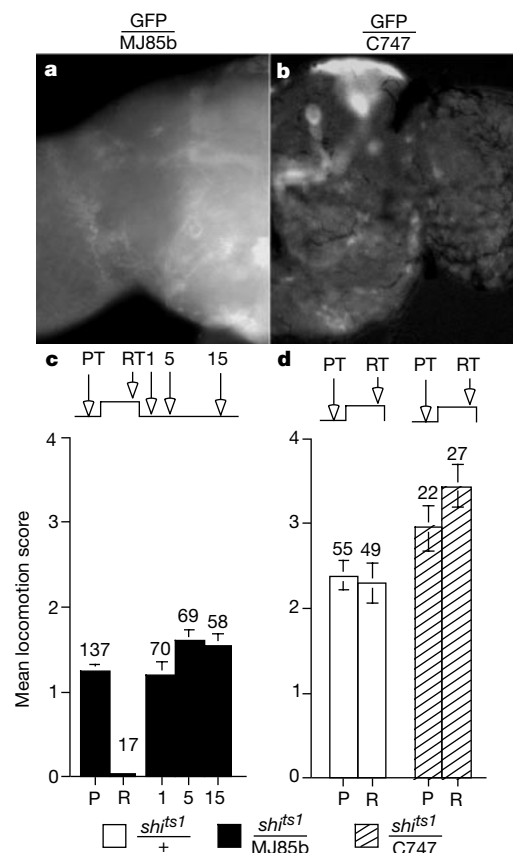


Figure 2 Disruption of neurotransmission produces paralysis when the *UAS-shi*^{ts1} transgene is expressed widely and has no effect when expression is restricted to the mushroom body. **a**, **b**, Whole mounts of adult brain from transgenic flies carrying a green fluorescent protein (GFP) reporter transgene (see Methods), show **(a)** widespread expression throughout the central brain (left hemisegment) when driven by the MJ85b GAL4 enhancer and **(b)** preferential expression in the mushroom body (right hemisegment) when driven by the C747 GAL4 enhancer. **c**, Locomotor activity in *shi*^{ts1}/MJ85b transgenic flies is abolished 30 min after shifting flies from permissive temperature (20 °C; PT) to restrictive temperature (30 °C; RT; $P < 0.001$). Recovery from paralysis is rapid. Locomotion appears normal again 1 min ($P = 0.166$), 5 min ($P = 0.222$) and 15 min ($P = 0.413$) after shifting transgenic flies back to PT (see Methods). **d**, Locomotion at PT or RT is similar for *shi*^{ts1}/+ control flies ($P = 0.546$) and *shi*^{ts1}/C747 transgenic flies ($P = 0.182$) after 30 min at RT.

procedure²⁵. Learning in this task was normal at permissive temperature in both *UAS-shi^{ts1}/C747* and *UAS-shi^{ts1}/C309* lines (Fig. 3a, permissive). When these transgenic flies were shifted to restrictive temperature 30 min before training, however, olfactory learning was abolished (Fig. 3a, restrictive). This temperature-dependent disruption of associative learning was not permanent. When *UAS-shi^{ts1}/C747* transgenic flies were shifted to restrictive temperature for 30 min and then back to permissive temperature for 30 min before training and testing, performance was normal (Fig. 3b, recovery).

Proper performance in learning experiments requires a combination of acquisition, storage and retrieval of memory. By limiting the temperature-dependent effects of *UAS-shi^{ts1}* to these three phases of memory formation, we were able to discern whether normal synaptic transmission in mushroom body neurons was required for each aspect of memory processing. We first asked whether synaptic transmission from the mushroom body was required during acquisition. Transgenic *UAS-shi^{ts1}/+*, *UAS-shi^{ts1}/C747* or *UAS-shi^{ts1}/C309* flies were shifted to restrictive temperature for 30 min, trained and then shifted back to permissive temperature immediately afterwards. These flies were then tested at permissive temperature after a 30-min retention period. Compared with control flies that were trained and tested at permissive temperature (Fig. 4a, control), performance was normal in these temperature-shifted flies (Fig. 4a, acquisition). This result indicates that synaptic transmission from mushroom body neurons is not required for acquisition of an odour-footshock association.

We next tested the requirement for mushroom body synaptic transmission during memory storage. Transgenic *UAS-shi^{ts1}/+*, *UAS-shi^{ts1}/C747* or *UAS-shi^{ts1}/C309* flies were trained at permissive temperature, shifted to restrictive temperature immediately after training and then shifted back to permissive temperature for 5 min before testing retention at 30 min. This treatment also did not adversely affect performance (Fig. 4a, storage), indicating that synaptic transmission in mushroom body neurons is not required for storage of early memory.

Finally, we assessed the requirement for mushroom body synaptic transmission during memory retrieval. Transgenic *UAS-shi^{ts1}/+*, *UAS-shi^{ts1}/C747* or *UAS-shi^{ts1}/C309* flies were trained at permissive temperature, shifted immediately afterwards to restrictive temperature and then tested for 30-min retention at the restrictive temperature. In contrast with the previous manipulations, memory retrieval appeared to be defective in these transgenic flies with disrupted synaptic transmission in mushroom body neurons (compare Fig. 4a, control and Fig. 4a, retrieval). This interpretation was complicated, however, by the fact that testing at restrictive temperature reduced 30-min retention scores of the *UAS-shi^{ts1}/+* control group (Fig. 4a). Thus, an alternative explanation for these data was that high temperature nonspecifically reduced performance in all genotypes.

We addressed this question by repeating the above experiment using a 5-min retention interval, for which general effects of restrictive temperature were not apparent (Fig. 3). Transgenic

UAS-shi^{ts1}/+, *UAS-shi^{ts1}/C747* or *UAS-shi^{ts1}/C309* flies were trained at permissive temperature, shifted to restrictive temperature immediately afterwards and then tested for 5-min retention at restrictive temperature. When compared with control flies that were trained and tested at permissive temperature (Fig. 4b, control), the performance of *UAS-shi^{ts1}/+* flies was not affected by testing at restrictive temperature (Fig. 4b, retrieval), as anticipated. In contrast, the performance of *UAS-shi^{ts1}/C747* or *UAS-shi^{ts1}/C309* flies was reduced significantly (Fig. 4b, retrieval). These results are similar to those in Fig. 3a, but were produced even when the transgenic flies were trained at permissive temperature but tested at restrictive temperature. As such, they confirm that disruption of synaptic transmission in the mushroom body blocks memory retrieval. We also used a reversal-retention protocol to dissociate general performance defects common to all the genotypes from a retrieval failure specific to the *UAS-shi^{ts1}/C747* and *UAS-shi^{ts1}/C309* flies. Using this protocol, we observed normal reversal learning at restrictive temperature in the *UAS-shi^{ts1}* control flies, but still observed defective retrieval in both experimental genotypes (see Supplementary Information).

Together, these data refine our understanding of mushroom body function during olfactory memory formation. Earlier studies in bees and flies have established that olfactory associative learning requires mushroom body neurons and cAMP signalling within them^{5,11,12}. The mushroom body is involved in associative mechanisms *per se* because neither pharmacological ablation of mushroom body neurons nor genetic perturbation of cAMP signalling within them affects the 'task-relevant' sensorimotor responses (olfactory

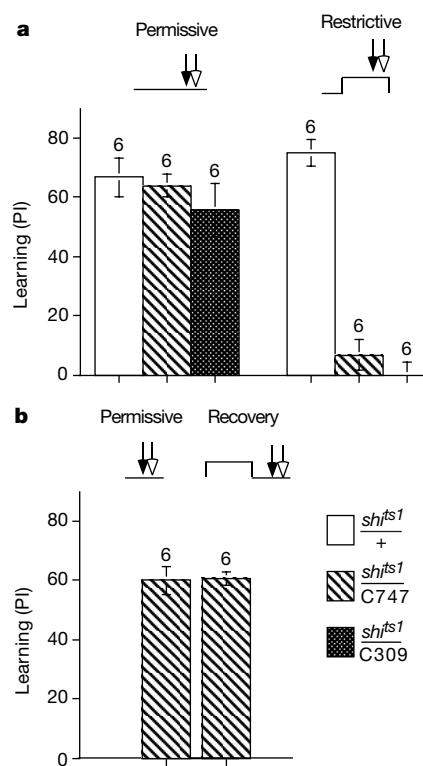


Figure 3 Disruption of neurotransmission in the mushroom body abolishes olfactory associative learning. **a**, When *shi^{ts1}/+*, *shi^{ts1}/C309* and *shi^{ts1}/C747* transgenic flies are trained (filled arrows) in a pavlovian olfactory conditioning assay (see Methods) and then tested (open arrows) immediately thereafter at PT (Permissive; see stimulus schedules above each panel), learning levels are high and similar among genotypes. When transgenic flies are shifted to RT (Restrictive) for 30 min and then trained and tested immediately thereafter, learning levels remain high in *shi^{ts1}/+* flies ($P = 0.355$) but are abolished in *shi^{ts1}/C309* ($P < 0.0001$) and *shi^{ts1}/C747* ($P < 0.0001$) flies. **b**, When *shi^{ts1}/C747* flies are shifted to RT for 30 min and then back to PT for a 30-min recovery before training and testing (Recovery), learning remains high ($P = 0.869$).

Table 1 Sensorimotor responses in *UAS-shi^{ts1}* flies

Genotype (temp.)	SR	OA (MCH)	OA (OCT)
<i>UAS-shi^{ts1}/+</i> (20 °C)	67 ± 9	63 ± 7	84 ± 4
<i>UAS-shi^{ts1}/C747</i> (20 °C)	56 ± 7	54 ± 11	63 ± 6
<i>UAS-shi^{ts1}/C309</i> (20 °C)	61 ± 13	59 ± 10	80 ± 7
<i>UAS-shi^{ts1}/+</i> (30 °C)	60 ± 10	37 ± 10	58 ± 4
<i>UAS-shi^{ts1}/C747</i> (30 °C)	54 ± 10	28 ± 5	37 ± 7
<i>UAS-shi^{ts1}/C309</i> (30 °C)	55 ± 9	22 ± 3	44 ± 2

Disruption of synaptic transmission in the mushroom body does not affect the 'task-relevant' sensorimotor responses (olfactory acuity, OA, and shock reactivity, SR) required for proper performance in pavlovian assays. The flies' 'task-relevant' abilities to sense and escape from the odours (olfactory acuity) or footshock (shock reactivity) were quantified in the T-maze (see Methods). Sensorimotor responses to footshock and both odours (OCT and MCH) were quantified at both restrictive and permissive temperatures. No significant differences were detected among genotypes at either temperature, although there was an overall effect of temperature on odour avoidance responses.

acuity and shock reactivity) necessary to perform this task. Connolly *et al.*¹¹ also established the spatial specificity of the enhancer-trap lines used here to drive *UAS-shi^{ts1}* expression in the mushroom body. Unlike these earlier studies, however, our use of the temperature-sensitive *UAS-shi^{ts1}* transgene permitted rapid and reversible disruption of synaptic transmission in otherwise intact mushroom body neurons.

In addition to its role in synaptic vesicle recycling, dynamin also participates in endocytotic protein trafficking of AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptors, a process that occurs on the timescale of several minutes and has been proposed to underlie the acquisition and/or early maintenance of long-term synaptic depression^{26,27}. An analogue of long-term depression may represent a cellular substrate of olfactory learning in *Drosophila*, but our disruption of dynamin function in mushroom body neurons did not affect either acquisition or storage of olfactory memory. Hence, our observations are more consistent with the theory that memory retrieval requires acute synaptic transmission from mushroom body axons, although the possibility that receptor internalization underlies memory retrieval cannot be ruled out formally.

These findings have several implications for our understanding of olfactory memory processing in *Drosophila*. First, they indicate that acquisition and storage of olfactory memory may both occur 'upstream' of synaptic transmission from mushroom body neurons. When combined with observations that cAMP signalling in the mushroom body is required for olfactory learning^{11,12}, our data

indicate that the hebbian processes underlying associative learning may reside within mushroom body neurons. Moreover, they are consistent with the idea that *rutabaga* adenylyl cyclase in mushroom body dendrites (calyces) may be the 'coincidence detector' for associations between the conditioned and unconditioned stimuli²⁸. This notion does not exclude the possibility that other hebbian processes upstream of the mushroom body, such as in the antennal lobes⁶, contribute to this form of learning. Finally, although mushroom body architecture suggests the presence of intrinsic feedback (Fig. 1)¹⁰, our data indicate that acquisition and memory storage during the first 30 min after training do not depend upon ongoing (reverberating) neuronal activity within the mushroom body. A recent study of *amnesiac* mutants is consistent with this view, although a later stage of memory formation (middle-term memory at 60-min retention) appears to depend on neural activity from (dorsal paired medial) neurons outside the mushroom body^{29,30}. In contrast, memory retrieval appears to be a neural process that requires transmission within (for example, feedback) and/or from mushroom body neurons. We propose that learning-driven changes in synaptic weight in mushroom body calyces contribute to a modification of output from the mushroom body during memory retrieval.

Understanding how newly acquired experiences are processed and coded will involve discovery of the molecular, cellular, network and behavioural properties underlying learning and memory formation. Rapidly reversible disruption of synaptic transmission in defined brain subregions permits a synthesis of temporally dynamic processes with identified anatomies. Computational models of cognition will derive from this integrative neurogenetic approach. □

Methods

Whole-mount expression of GFP

To examine nervous systems of *Drosophila melanogaster* expressing green fluorescent protein (GFP), brains were dissected in 4% paraformaldehyde in PBS. They were fixed by being maintained in this solution for more than 30 min. They were then mounted in Vectashield (Vector Labs) and examined under a Zeiss fluorescent microscope with a far-blue (FITC) filter. *UAS-shi^{ts1}/MJ85b*, *UAS-shi^{ts1}/C747*, *UAS-shi^{ts1}/C309* and *UAS-shi^{ts1}/+* flies were generated by crossing *MJ85b/MJ85b*²³, *C747/C747* and *C309/C309* homozygotes or *w¹¹¹⁸* (isoCJ1) inbred wild-type flies with *UAS-shi^{ts1}/UAS-shi^{ts1}*; *UAS-shi^{ts1}/UAS-shi^{ts1}* double homozygotes⁷. *UAS-GFP/MJ85b* and *UAS-GFP/C747* flies were generated by crossing *UAS-GFP* homozygotes with *MJ85b* or *C747* homozygotes, respectively. *C747* and *C309* enhancer trap lines drive preferential expression in the mushroom body, but also yield some expression elsewhere. Hence we have used these two independent lines throughout to focus on mushroom body neurons, the common anatomical site of expression in both lines.

Behavioural analyses

Locomotor activity was measured using countercurrent phototaxis apparatus²⁴ in the absence of light. Under these conditions, tapping flies into the start tube between trials produces an escape response, generally referred to as locomotor reactivity, which decays with time until only spontaneous locomotor activity is expressed. Five runs were done to separate flies, fractionating them into six groups based on the number of times a fly ran into the distal tube. The most active flies ran into distal tubes five times and received a score of five; the least active flies stayed in the original start tube throughout the experiment and received a score of zero. Countercurrent data are binomially distributed. Hence, these data were subjected to an arcsin square-root transformation before subjecting them to analysis of variance. To maintain an experiment-wise error rate of $\alpha = 0.05$ for data in Fig. 2c and d, the adjusted error rate (α') for the six subsequent planned pairwise comparisons was 0.009.

Olfactory associative learning was quantified by subjecting 2–3-day-old adult flies to a pavlovian conditioning procedure²⁵. Groups of about 100 flies received one training session, during which they were exposed sequentially to one odour (conditioned stimulus, CS+) paired with footshock and then a second odour (CS-) without footshock. Conditioned odour avoidance was tested immediately or 5 min or 30 min after training. During the test trial, flies were exposed simultaneously to the CS+ and CS- in a T-maze. After 2 min, flies were trapped in either T-maze arm, anaesthetized and counted. From this distribution, a performance index (PI) was calculated, so that a 50:50 distribution (no memory) yielded a PI of zero and a 0:100 distribution away from the CS+ yielded a PI of 100.

Olfactory acuity was quantified by exposing naive flies to odour (at concentration used for pavlovian training) versus air in the T-maze during a 2-min test trial. Shock reactivity was quantified by placing grids in each arm of the T-maze and then by exposing naive flies to shock versus no shock (at intensity used for pavlovian training) during a 2-min test trial. For both olfactory acuity and shock reactivity, PIs were calculated as above. The

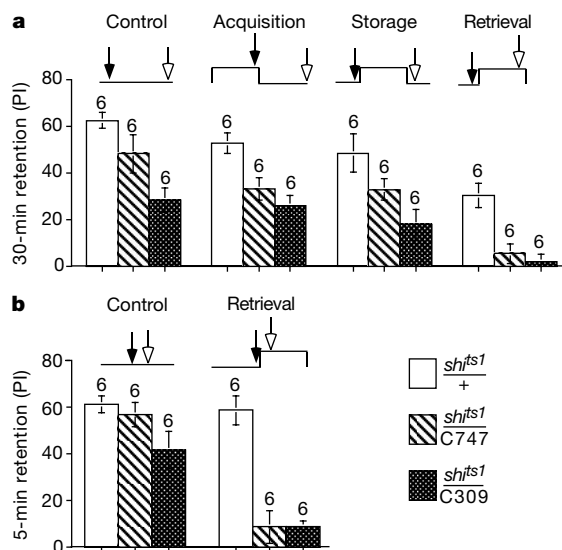


Figure 4 Disruption of neurotransmission in the mushroom body abolishes retrieval but not acquisition or storage of olfactory associative memory. **a**, Control: when *shi^{ts1}/+*, *shi^{ts1}/C309* and *shi^{ts1}/C747* transgenic flies are trained and tested 30 min later at PT, all three groups show learning, although retention levels are lower in *shi^{ts1}/C309* ($P < 0.001$) but not in *shi^{ts1}/C747* flies ($P = 0.067$) than in *shi^{ts1}/+* controls. Acquisition: when flies are shifted to RT for 30 min, trained and then shifted back to PT for 30 min before testing, retention in all three genotypes is similar to that at PT ($P = 0.212$, 0.757 and 0.056 for *shi^{ts1}/+*, *shi^{ts1}/C309* and *shi^{ts1}/C747*, respectively). Storage: when flies are trained at PT, shifted to RT for a 30-min retention interval and then shifted back to PT for 5 min before testing, performance levels are again similar to those at PT ($P = 0.072$, 0.176 and 0.049 for *shi^{ts1}/+*, *shi^{ts1}/C309* and *shi^{ts1}/C747*, respectively). Retrieval: when transgenic flies are trained at PT, shifted to RT immediately and tested for 30-min retention, memory remains in *shi^{ts1}/+* control flies but is near zero in *shi^{ts1}/C309* ($P = 0.607$) and *shi^{ts1}/C747* flies ($P = 0.243$). **b**, Control: when flies are trained and tested 5 min later at PT, all three genotypes display robust performance. Retrieval: when flies are trained at PT, shifted to RT immediately and then tested for 5-min retention, memory scores are unaffected for *shi^{ts1}/+* control flies ($P = 0.77$) but are significantly reduced for *shi^{ts1}/C747* ($P < 0.001$) and *shi^{ts1}/C309* flies ($P < 0.001$).

adjusted error rate (α') for the six subsequent planned pairwise comparisons was 0.009. $n = 8$ PIs for each group.

PIs are distributed normally. Hence, these data were analysed by analysis of variance with subsequent pairwise comparisons adjusted to maintain an experiment-wise error rate of $\alpha = 0.05$. In Fig. 3, the error rate for the one subsequent planned comparison was 0.05. In Fig. 4, to maintain an experiment-wise error rate of $\alpha = 0.05$, the adjusted error rates (α') were $P = 0.004$ and 0.01 , respectively, for the 13 subsequent planned pairwise comparisons in Fig. 4a and for the 5 subsequent planned comparisons in Fig. 4b. The number of PIs included in each group mean is listed above the corresponding bar in all figures.

Received 25 January; accepted 28 March 2001.

- Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neuropsychiat. Clin. Neurosci.* **12**, 103–113 (2000).
- Milner, B., Squire, L. R. & Kandel, E. R. Cognitive neuroscience and the study of memory. *Neuron* **20**, 445–468 (1998).
- Mayford, M. *et al.* Control of memory formation through regulated expression of a CaMKII transgene. *Science* **274**, 1678–1683 (1996).
- Shimizu, E., Tang, Y. P., Rampon, C. & Tsien, J. Z. NMDA receptor-dependent synaptic reinforcement as a crucial process for memory consolidation. *Science* **290**, 1170–1174 (2000).
- de Belle, J. S. & Heisenberg, M. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–695 (1994).
- Hammer, M. & Menzel, R. Multiple sites of associative odor learning as revealed by local brain microinjections of octopamine in honeybees. *Learn. Mem.* **5**, 146–156 (1998).
- Kitamoto, T. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive *shibire* allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).
- Rybak, J. & Menzel, R. Anatomy of the mushroom bodies in the honey bee brain: the neuronal connections of the α -lobe. *J. Comp. Neurol.* **334**, 444–465 (1993).
- Strausfeld, N. J. & Li, Y. Organization of olfactory and multimodal afferent neurons supplying the calyx and pedunculus of the cockroach mushroom bodies. *J. Comp. Neurol.* **409**, 603–625 (1999).
- Ito, K. *et al.* The organization of extrinsic neurons and their implications in the functional roles of the mushroom bodies in *Drosophila melanogaster* Meigen. *Learn. Mem.* **5**, 52–77 (1998).
- Connolly, J. B. *et al.* Associative learning disrupted by impaired Gs signaling in *Drosophila* mushroom bodies. *Science* **274**, 2104–2106 (1996).
- Zars, T., Fischer, M., Schulz, R. & Heisenberg, M. Localization of a short-term memory in *Drosophila*. *Science* **288**, 672–675 (2000).
- Poodry, C. A., Hall, L. & Suzuki, D. T. Developmental properties of *Shibire*: a pleiotropic mutation affecting larval and adult locomotion and development. *Dev. Biol.* **32**, 373–386 (1973).
- Poodry, C. A. & Edgar, L. Reversible alteration in the neuromuscular junctions of *Drosophila melanogaster* bearing a temperature-sensitive mutation, *shibire*. *J. Cell Biol.* **81**, 520–527 (1979).
- Ramaswami, M., Rao, S., van der Blik, A., Kelly, R. B. & Krishnan, K. S. Genetic studies on dynam function in *Drosophila*. *J. Neurogenet.* **9**, 73–87 (1993).
- Koenig, J. H. & Ikeda, K. Disappearance and reformation of synaptic vesicle membrane upon transmitter release observed under reversible blockage of membrane retrieval. *J. Neurosci.* **9**, 3844–3860 (1989).
- Chen, M. S. *et al.* Multiple forms of dynamin are encoded by *shibire*, a *Drosophila* gene involved in endocytosis. *Nature* **351**, 583–586 (1991).
- van der Blik, A. M. & Meyerowitz, E. M. Dynamin-like protein encoded by the *Drosophila shibire* gene associated with vesicular traffic. *Nature* **351**, 411–414 (1991).
- Damke, H., Baba, T., Warnock, D. E. & Schmid, S. L. Induction of mutant dynamin specifically blocks endocytic coated vesicle formation. *J. Cell Biol.* **127**, 915–934 (1994).
- Grant, D., Unadkat, S., Katzen, A., Krishnan, K. S. & Ramaswami, M. Probable mechanisms underlying interallelic complementation and temperature-sensitivity of mutations at the *shibire* locus of *Drosophila melanogaster*. *Genetics* **149**, 1019–1030 (1998).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Yang, M. Y., Armstrong, J. D., Vilinsky, I., Strausfeld, N. J. & Kaiser, K. Subdivision of the *Drosophila* mushroom bodies by enhancer-trap expression patterns. *Neuron* **15**, 45–54 (1995).
- Joiner, M. A. & Griffith, L. C. Mapping of the anatomical circuit of CaM kinase-dependent courtship conditioning in *Drosophila*. *Learn. Mem.* **6**, 177–192 (1999).
- Benzer, S. Behavioral mutants of *Drosophila* isolated by countercurrent distribution. *Proc. Natl Acad. Sci. USA* **58**, 1112–1119 (1967).
- Tully, T., Preat, T., Boynton, S. C. & Del Vecchio, M. Genetic dissection of consolidated memory in *Drosophila*. *Cell* **79**, 35–47 (1994).
- Wang, Y. T. & Linden, D. J. Expression of cerebellar long-term depression requires postsynaptic clathrin-mediated endocytosis. *Neuron* **25**, 635–647 (2000).
- Carroll, R. C. *et al.* Dynamin-dependent endocytosis of ionotropic glutamate receptors. *Proc. Natl Acad. Sci. USA* **96**, 14112–14117 (1999).
- Dudai, Y., Corfas, G. & Hazy, S. What is the possible contribution of Ca^{2+} -stimulated adenylate cyclase to acquisition, consolidation and retention of an associative olfactory memory in *Drosophila*? *J. Comp. Physiol. A* **162**, 101–109 (1988).
- Waddell, S., Armstrong, J. D., Kitamoto, T., Kaiser, K. & Quinn, W. G. The amnesiac gene product is expressed in two neurons in the *Drosophila* brain that are critical for memory. *Cell* **103**, 805–813 (2000).
- Dubnau, J. & Tully, T. Functional anatomy: from molecule to memory. *Curr. Biol.* **11**, R240–R243 (2001).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank H. Cline, E. Drier, J. Huang, M. Regulski and K. Svoboda for comments on the manuscript. This work was supported by the NIH (J.T.D. and T.T.) and the NSF (T.K.).

Correspondence and requests for materials should be addressed to J.D. (e-mail: dubnau@cshl.org).

Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus

Michael A. Cowley^{*†}, James L. Smart^{*†}, Marcelo Rubinstein[‡], Marcelo G. Cerdán[‡], Sabrina Diano[§], Tamas L. Horvath^{§||}, Roger D. Cone^{*} & Malcolm J. Low^{*}

^{*} The Vollum Institute, Oregon Health Sciences University, Portland, Oregon 97201-3098, USA

[‡] Instituto de Investigaciones en Ingeniería Genética y Biología Molecular, CONICET and Department of Biology, School of Sciences, University of Buenos Aires 1428, Argentina

[§] Reproductive Neurosciences Unit, Department of Obstetrics and Gynecology and ^{||} Department of Neurology, Yale Medical School, New Haven, Connecticut 06520, USA

[†] These authors contributed equally to this work

The administration of leptin¹ to leptin-deficient humans, and the analogous *Lep^{ob}/Lep^{ob}* mice, effectively reduces hyperphagia and obesity^{2,3}. But common obesity is associated with elevated leptin, which suggests that obese humans are resistant to this adipocyte hormone. In addition to regulating long-term energy balance, leptin also rapidly affects neuronal activity^{4–6}. Proopiomelanocortin (POMC) and neuropeptide-Y types of neurons in the arcuate nucleus of the hypothalamus⁷ are both principal sites of leptin receptor expression and the source of potent neuropeptide modulators, melanocortins and neuropeptide Y, which exert opposing effects on feeding and metabolism^{8,9}. These neurons are therefore ideal for characterizing leptin action and the mechanism of leptin resistance; however, their diffuse distribution makes them difficult to study. Here we report electrophysiological recordings on POMC neurons, which we identified by targeted expression of green fluorescent protein in transgenic mice. Leptin increases the frequency of action potentials in the anorexigenic POMC neurons by two mechanisms: depolarization through a nonspecific cation channel; and reduced inhibition by local orexigenic neuropeptide-Y/GABA (γ -aminobutyric acid) neurons. Furthermore, we show that melanocortin peptides have an autoinhibitory effect on this circuit. On the basis of our results, we propose an integrated model of leptin action and neuronal architecture in the arcuate nucleus of the hypothalamus.

Previous studies suggest that leptin does not have equal effects on all neuronal subtypes. Acute leptin treatment presumably activates POMC, but not neuropeptide Y (NPY) neurons in the arcuate nucleus of the hypothalamus (ARC), because c-Fos protein is increased only in the former population and Socs3 messenger RNA is increased in both¹⁰. Furthermore, a population of ARC neurons seems to be inhibited directly by leptin, but the peptide phenotype of these neurons has not been directly established^{11–14}. In addition to leptin receptors, both POMC and NPY neurons express a receptor, MC3-R, for POMC-derived melanocortin peptides¹⁵. The physiological role of this receptor is not well understood, although MC3-R null mice have increased adiposity compared with wild-type mice¹⁶.

To test the hypothesis that leptin selectively activates POMC neurons, we first generated a strain of transgenic mice expressing green fluorescent protein (EGFP; Clontech) under the transcriptional control of mouse *Pomc* genomic sequences, including a region located between –13 kilobases (kb) and –2 kb that is required for accurate neuronal expression¹⁷ (Fig. 1a). Bright green fluorescence (509 nm) was seen in the two central nervous system regions where POMC is produced: the ARC and the nucleus of the solitary tract (data not shown).

Under ultraviolet (450–480 nm) excitation, POMC neurons were clearly distinguished from adjacent, non-fluorescent neurons (Fig. 1b) visualized under infrared optics. Double immunofluorescence showed that there was more than 99% cellular colocalization of EGFP and POMC peptides in the ARC (Fig. 1c). There was close apposition of both tyrosine-hydroxylase- and NPY-stained terminals on EGFP-expressing POMC neurons, but no evidence of colocalization of the tyrosine hydroxylase or NPY immunoreactivity with EGFP (data not shown). Total fluorescent cell counts of coronal hypothalamic sections determined that there were

$3,148 \pm 62$ (mean $\pm$ s.e.m.; $n = 3$) POMC-EGFP neurons distributed through the whole ARC¹⁸ (Fig. 1d). POMC neurons are located both medially and ventrally in mouse ARC, in contrast to a predominantly lateral position in rat ARC.

POMC-EGFP neurons in hypothalamic slices had a resting membrane potential of -40 to -45 mV and exhibited frequent spontaneous action potentials. The non-selective opioid agonist met-enkephalin (Met-Enk, $30 \mu\text{M}$; Sigma) caused a rapid (35 – 40 s), reversible hyperpolarization (10 – 20 mV) of the membrane potential of POMC cells ($n = 10$) and prevented spontaneous action potential generation (Fig. 2a). In normal (2.5 mM K^+) Krebs solution, the reversal potential of the inwardly rectifying opioid current was about -90 mV, whereas in 6.5 mM K^+ Krebs solution the reversal potential shifted to about -60 mV ($n = 3$; Fig. 2b). The μ -opioid receptor (MOP-R) antagonist CTAP ($1 \mu\text{M}$; Phoenix Pharmaceuticals) completely prevented the current induced by Met-Enk in POMC cells ($n = 3$; Fig. 2c). These characteristics indicate that the opioid current was caused by both activation of MOP-R and increased ion conductance through G-protein-coupled, inwardly rectifying potassium channels¹⁹. The similarity of opioid responses in EGFP-labelled POMC neurons to those of guinea-pig¹⁹ or

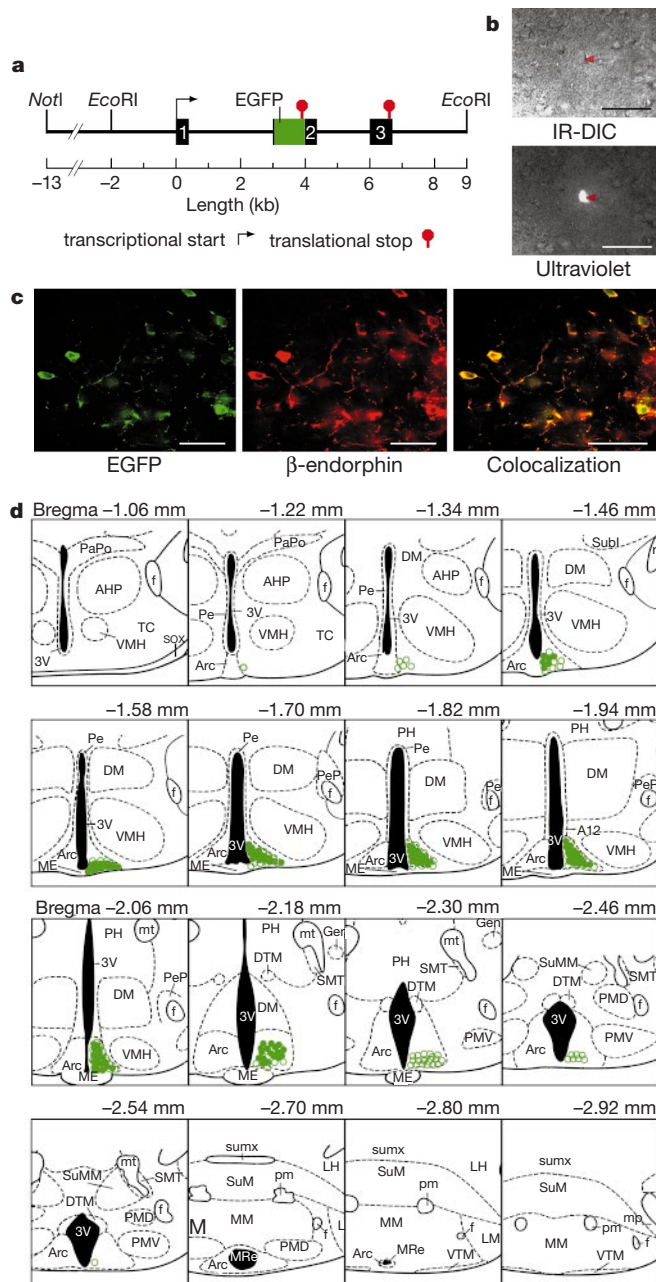


Figure 1 Generation of transgenic mice expressing EGFP in ARC POMC neurons. **a**, Structure of the POMC-EGFP transgene. **b**, Identification of a single POMC neuron (red arrowhead on recording electrode tip) by infrared-differential interference contrast (IR-DIC) microscopy (upper) and EGFP fluorescence (lower) in a living ARC slice before electrophysiological recordings. **c**, Colocalization (yellow) of EGFP (green) and β -endorphin immunoreactivity (red) in ARC POMC neurons. Scale bars, $50 \mu\text{m}$ (**b**, **c**). **d**, Distribution of EGFP-positive neuronal soma throughout the ARC nucleus. Green open circles = 5 cells, green filled circles = 10 cells.

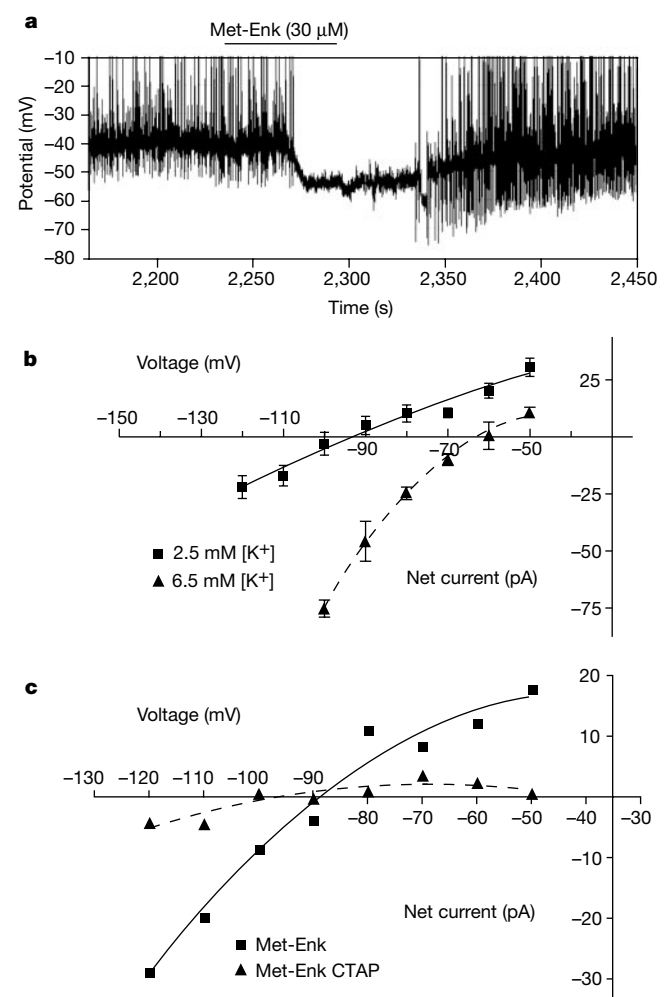


Figure 2 Activation of MOP-Rs hyperpolarizes the EGFP-labelled POMC neurons by opening G-protein-coupled inwardly rectifying potassium channels. **a**, Met-enkephalin (Met-Enk) hyperpolarizes POMC neurons and inhibits all action potentials. Horizontal bar indicates the time when $30 \mu\text{M}$ Met-Enk was bath applied to the slice. **b**, Met-Enk current and reversal potential is shifted by extracellular K^+ concentration. **c**, Met-Enk activates MOP-Rs on POMC neurons. A Met-Enk ($30 \mu\text{M}$) current was observed, and the MOP-R-specific antagonist CTAP ($1 \mu\text{M}$) was applied for 1 min. After CTAP, Met-Enk elicited no current. Figure is representative of three experiments.

mouse²⁰ POMC cells, identified by post-recording immunohistochemistry, suggests that expression of the EGFP transgene does not compromise expression of receptors nor their coupling to second messenger systems in POMC neurons.

We next determined the direct effects of leptin on identified POMC cells in slice preparations. Leptin (0.1–100 nM) depolarized 72 out of 77 POMC cells by 3–30 mV (Fig. 3a; depolarization at 100 nM leptin, 9.7 ± 1.2 mV (mean $\pm$ s.e.m.; $n = 45$)) in 2–10 min, in a concentration-responsive manner (Fig. 3b). There were two components to the depolarization and neither were fully reversible within 40 min. First, the depolarization was due to a small inward current that reversed at about -20 mV (Fig. 3c), suggesting the involvement of a nonspecific cation channel²¹. Second, leptin treatment decreased the GABA-mediated tone onto POMC cells. GABA-mediated inhibitory postsynaptic currents (IPSCs) were observed in POMC cells, and leptin (100 nM) decreased their frequency by 25% (Fig. 3d) in 5 out of 15 cells, indicating that it may act presynaptically to reduce GABA release (leptin had no effect on IPSCs in 10 out of 15 POMC neurons).

The effect on IPSC frequency occurred with a similar lag to the effect on membrane potential. Thus, leptin not only directly depolarizes POMC neurons but also acts at GABA-secreting nerve terminals to reduce the release of GABA onto POMC neurons, allowing them to adopt a more depolarized resting potential. The consistent depolarization of POMC cells by leptin was specific because leptin had no effect on 5 out of 13 adjacent non-fluorescent cells tested (Fig. 3e), whereas it hyperpolarized five (Fig. 3f) and depolarized three other non-POMC neurons in the ARC (data not shown). The electrophysiological effects of leptin reported here are consistent with leptin's biological actions—leptin rapidly causes release of α -melanocyte-stimulating hormone (α -MSH) from rat hypothalamus⁴, presumably by activating POMC neurons.

Previous reports of neuronal hyperpolarization by leptin^{11,12} and the colocalization of GABA and NPY²² in subpopulations of ARC neurons led us to speculate that leptin hyperpolarizes NPY/GABA cells that directly innervate POMC neurons, and thus reduces GABA-mediated drive onto POMC cells. Both leptin and NPY Y2 receptors are expressed on NPY neurons in the ARC^{7,23}. Furthermore, activation of Y2 receptors inhibits NPY release from NPY neurons²⁴, and presumably would also diminish GABA release from NPY/GABA terminals. This provided us with an alternative pharmacological approach, independent of leptin, to test the proposed innervation of POMC neurons by GABA-secreting NPY neurons.

Indeed, NPY (100 nM; Bachem) decreased the frequency of GABA-mediated IPSCs by 55% in 3 min in all 12 POMC cells tested (Fig. 4a). Both NPY and leptin still inhibited IPSCs in the presence of tetrodotoxin (TTX; 6/6 and 3/5 cells, respectively), indicating that some of the inhibition of IPSCs was occurring through direct effects at presynaptic nerve terminals. POMC neurons express the NPY Y1 receptor²³, and NPY also hyperpolarized all POMC neurons tested by an average of 9 ± 6 mV ($n = 3$; data not shown).

To confirm the origin of GABA-mediated innervation on POMC neurons from NPY/GABA terminals, we also tested the effect of the highly selective MC3-R agonist D-Trp⁸- γ MSH (ref. 25) on local GABA release. D-Trp⁸- γ MSH (7 nM) increased the frequency of GABA-mediated IPSCs ($280 \pm 90\%$) recorded from three out of four POMC neurons (Fig. 4b). It had no effect on one cell. The positive effect of MC3-R activation, together with the negative effects of NPY and leptin, shows the dynamic range of the NPY/GABA synapse onto POMC neurons and points to the important role of this synapse in modulating signal flow in the ARC. D-Trp⁸- γ MSH (7 nM) also hyperpolarized (-5.5 ± 2.4 mV) 9 out of 15 POMC neurons tested, and decreased the frequency of action potentials (Fig. 4c); the remaining cells showed no significant response to D-Trp⁸- γ MSH. We could not determine whether these effects were entirely due to increased GABA release onto the POMC

cells, or whether there was an additional postsynaptic action of D-Trp⁸- γ MSH on POMC neurons, roughly half of which also express the MC3-R¹⁵. Thus, MC3-R acts in a similar autoreceptor manner to MOP-Rs on POMC neurons, diminishing POMC neuronal activity in response to elevated POMC peptides.

To determine that the IPSCs in POMC neurons were due to local innervation by NPY/GABA cells, we carried out multi-label immuno-

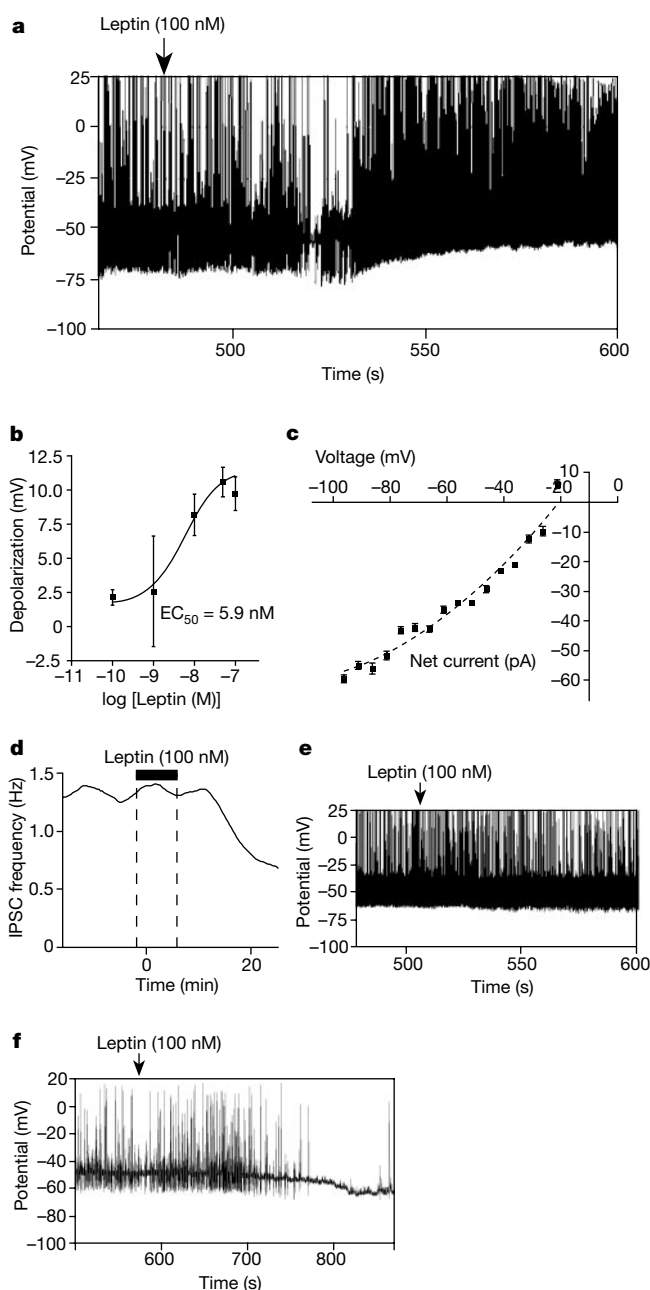


Figure 3 Leptin depolarizes POMC neurons through a nonspecific cation channel, and decreases the GABA-mediated tone onto POMC cells. **a**, Leptin depolarizes POMC neurons and increases the frequency of action potentials within 1–10 min of addition. Figure is representative of recordings made from 77 POMC neurons. **b**, Leptin causes a concentration-dependent depolarization of POMC cells. Depolarization by leptin was determined at 0.1, 1, 10, 50 and 100 nM (effector concentration at half-maximum response, $EC_{50} = 5.9$ nM) in 8, 7, 9, 3 and 45 cells, respectively. **c**, Leptin depolarizes POMC cells by activating a nonspecific cation current. Figure is representative of the response in ten cells. **d**, Leptin decreases the frequency of IPSCs in POMC cells. Figure is an example of five cells in which leptin (100 nM) decreased the frequency of IPSCs. **e**, Leptin had no effect on five adjacent non-fluorescent ARC neurons. **f**, Leptin hyperpolarized five non-fluorescent ARC neurons.

histochemistry using light and electron microscopy. Although independent NPY²⁷ and GABA²⁷ innervation of POMC cells has been reported, colocalization of NPY and GABA in nerve terminals forming synapses onto POMC cells has not been shown. Similar to the rat²⁶, a dense innervation of POMC cells by NPY axon terminals was detected in the mouse (Fig. 4d). Electron microscopy confirmed the co-expression of NPY and GABA in axon terminals, and showed

that these boutons established synapses on the perikarya of all 15 ARC POMC neurons analysed (Fig. 4e).

These observations allow us to propose a detailed model of regulation of this circuit including: dual mechanisms of leptin action in the ARC, interactions between NPY/GABA and POMC neurons, and autoregulatory feedback from opioid and melanocortin peptides as well as NPY (Fig. 4f). In this model, leptin directly depolarizes the POMC neurons while simultaneously hyperpolarizing the somata of NPY/GABA neurons, and diminishes release from NPY/GABA terminals. This diminished GABA release disinhibits the POMC neurons. Together, the direct and indirect effects of leptin result in an activation of POMC neurons and an increased frequency of action potentials. In addition, both POMC and NPY neurons express autoreceptors for some of their respective neuropeptide products (β -endorphin or α -MSH, and NPY, respectively) and activation of these autoreceptors may provide ultrashort feedback loops that further modulate the effects of leptin on POMC neurons. The effects of NPY and melanocortin agonists on IPSC frequency are similar to the effect of NPY on IPSCs in the paraventricular nucleus of the hypothalamus²⁸.

Thus, there seem to be two classes of neurons accounting for leptin sensitivity in the brain: those activated (depolarized) to release anorexigenic peptides; and those inhibited (hyperpolarized) with a consequent reduction in release of orexigenic peptides. Identifying viable POMC neurons with EGFP provides an invaluable tool for determining complex actions on these important neurons. Furthermore, an understanding of the acute effects of leptin provides a model system for studying the mechanisms underlying the development of leptin resistance²⁹. □

Methods

Generation of POMC-EGFP mice

The *EGFP* cassette contains its own Kozak consensus translation initiation site, along with *SV40* polyadenylation signals downstream of the *EGFP* coding sequences, which directs proper processing of the 3' end of the *EGFP* mRNA. We introduced the *EGFP* cassette by standard techniques into the 5' untranslated region of exon 2 of a mouse *Pomc* genomic clone containing 13 kb of 5' and 2 kb of 3' flanking sequences¹⁷. The transgene was microinjected into pronuclei of one-cell-stage embryos of C57BL/6J mice (Jackson Laboratories) as described¹⁷. One founder was generated and bred to wild-type C57BL/6J to produce N₁ hemizygous mice. In addition, we also generated N₂ and subsequent generations of mice homozygous for the transgene. The mice are fertile and have normal growth and development.

Immunofluorescence and GFP colocalization

Anaesthetized mice were perfused transcardially with 4% paraformaldehyde, and free-floating brain sections were prepared with a vibratome. We processed sections for immunofluorescence and colocalization of GFP fluorescence using standard techniques. Primary antisera and their final dilutions were rabbit anti- β -endorphin, 1:2500 (v/v); rabbit anti-NPY, 1:25,000 (v/v) (Alanax Corp.); rabbit anti-ACTH, 1:2000 (v/v); and mouse anti-TH, 1:1000 (v/v) (Incstar). After rinsing, sections were incubated with 10 mg ml⁻¹ biotinylated horse anti-mouse/rabbit immunoglobulin- γ (IgG) (Vector Laboratories) followed by Cy-3 conjugated streptavidin, 1:500 (v/v) (Jackson Immuno-research Laboratories). Photomicrographs were taken on an Axioscop (Zeiss) using fluorescein isothiocyanate (FITC) and rhodamine isothiocyanate (RITC) filter sets (Chroma Technology Corp.).

Electrophysiology

We cut 200- μ m thick coronal slices from the ARC of 4-week-old male POMC-EGFP mice. Slices were maintained at 35 °C in Krebs solution ((in mM): 126 NaCl, 2.5 KCl, 1.2 MgCl₂, 2.4 CaCl₂·2H₂O, 1.2 NaH₂PO₄·H₂O, 21.4 NaHCO₃, 11.1 glucose), and were saturated with 95% O₂, 5% CO₂ for 1 h before recordings. Recordings were made in Krebs solution at 35 °C. Slices were visualized on an Axioscop FS2 (Zeiss) through standard infrared optics and using epifluorescence through a FITC filter set (see Fig. 1c). Whole-cell recordings were made from fluorescent neurons using an Axopatch 1D amplifier (Axon Instruments) and Clampex 7 (Axon Instruments).

We determined resting membrane potentials using an event-detection protocol on a PowerLab system (AD Instruments) to average expanded traces of the membrane potential. Drugs were applied to the bath over the times indicated. The resting membrane potential was stable for up to 1 h in cells treated with Krebs solution alone (data not shown). *I*-*V* relationships for the Met-Enk currents were established using a step protocol (-60 mV holding potential, sequentially pulsed (40 ms) from -120 to -50 mV; cells were returned to -60 mV for 2 s between voltage steps). The protocol was repeated after addition of Met-Enk. The net current was the difference between the two *I*-*V* relation-

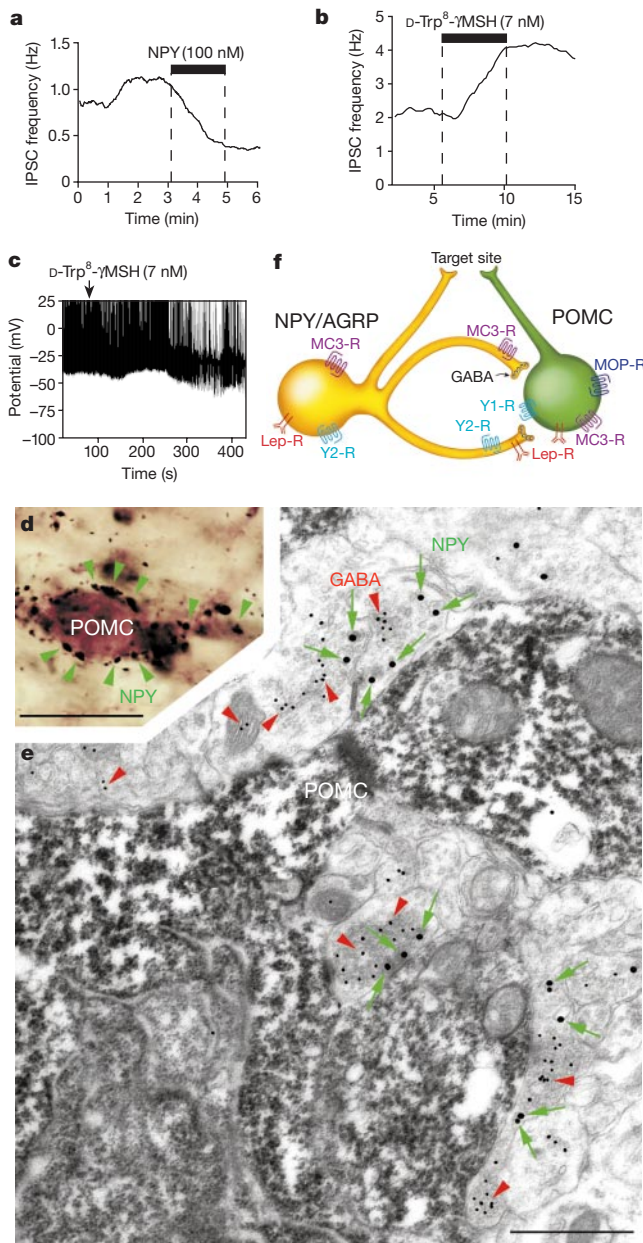


Figure 4 The GABA-mediated inputs to POMC cells are from NPY neurons that co-express GABA. **a**, NPY decreases the frequency of mini-IPSCs in POMC neurons. **b**, D-Trp⁸- γ MSH (7nM)—a dose that selectively activates MC3-R—increases the frequency of GABA-mediated IPSCs in POMC neurons. **c**, D-Trp⁸- γ MSH hyperpolarizes POMC neurons. Panels **a**–**c** are representative results. **d**, Expression of NPY in nerve terminals adjacent to POMC neurons in the ARC. NPY nerve terminals (black, green arrowheads); POMC neuronal soma (brown). Scale bar, 10 μ m. **e**, Expression of GABA and NPY in nerve terminals synapsing onto POMC neurons in the ARC. GABA immunoreactivity (10-nm gold particles, red arrowheads) and NPY immunoreactivity (25-nm gold particles, green arrows) are in separate vesicle populations colocalized in synaptic boutons that make direct contact with the soma of POMC neurons (DAB contrasted with uranyl acetate and lead citrate, diffuse black in cytoplasm). Scale bar, 1 μ m. **f**, Model of leptin regulation of NPY/GABA and POMC neurons in the ARC (see text).

ships. This protocol was repeated in Krebs solution with 6.5 mM K⁺. To identify the postsynaptic leptin current, *I*-*V* relationships were performed similarly with slow voltage ramps (5 mV s⁻¹ from -100 to -20 mV) before and 10 min after adding leptin (100 nM).

GABA-mediated IPSCs were recorded using a CsCl internal electrode solution ((in mM): 140 CsCl, 10 HEPES, 5 MgCl₂, 1 BAPTA, 5 Mg-ATP, 0.3 Na-GTP). Both mini IPSCs and large amplitude (presumably multisynaptic) IPSCs were observed in the untreated slices. TTX (1 μM) abolished large IPSCs. We acquired data before and after drug addition at a -50-mV holding potential in 2-s sweeps every 4 s for the times indicated in the figures. Mini-postsynaptic currents were analysed using Axograph 4 (Axon Instruments). IPSCs and excitatory postsynaptic currents (EPSCs) were distinguished on the basis of their decay constants; in addition, picrotoxin (100 μM) blocked all IPSCs. POMC neurons receive a low EPSC tone, and the frequency was not modulated by any of the treatments described here.

Immunostaining for light and electron microscopy

We carried out double immunocytochemistry for NPY and POMC using different colour diaminobenzidine (DAB) chromogens on fixed mouse hypothalami according to published protocols²⁷. For electron microscopy, pre-embedding immunostaining for β-endorphin was done with an ABC Elite kit (Vector Laboratories) and a DAB reaction, followed by post-embedding labelling of GABA and NPY using rabbit anti-GABA, 1:1000 (v/v), and gold-conjugated (10-nm) goat anti-rabbit IgG or sheep anti-NPY and gold-conjugated (25-nm) goat anti-sheep IgG. Sections were contrasted with saturated uranyl acetate (10 min) and lead citrate (20–30 s), and examined using a Philips CM-10 electron microscope.

Received 19 December 2000; accepted 12 March 2001.

1. Zhang, Y. *et al.* Positional cloning of the mouse obese gene and its human homologue. *Nature* **372**, 425–432 (1994); erratum *ibid.* **374**, 479 (1995).
2. Farooqi, I. S. *et al.* Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N. Engl. J. Med.* **341**, 879–884 (1999).
3. Campfield, L. A., Smith, F. J., Guise, Y., Devos, R. & Burn, P. Recombinant mouse OB protein: evidence for a peripheral signal linking adiposity and central neural networks. *Science* **269**, 546–549 (1995).
4. Kim, M. S. *et al.* The central melanocortin system affects the hypothalamo-pituitary thyroid axis and may mediate the effect of leptin. *J. Clin. Invest.* **105**, 1005–1011 (2000).
5. Haynes, W. G., Morgan, D. A., Walsh, S. A., Mark, A. L. & Sivit, W. I. Receptor-mediated regional sympathetic nerve activation by leptin. *J. Clin. Invest.* **100**, 270–278 (1997).
6. Haynes, W. G., Morgan, D. A., Djalali, A., Sivit, W. I. & Mark, A. L. Interactions between the melanocortin system and leptin in control of sympathetic nerve traffic. *Hypertension* **33**, 542–547 (1999).
7. Hakansson, M. L., Brown, H., Ghilardi, N., Skoda, R. C. & Meister, B. Leptin receptor immunoreactivity in chemically defined target neurons of the hypothalamus. *J. Neurosci.* **18**, 559–572 (1998).
8. Kalra, S. P. *et al.* Interacting appetite-regulating pathways in the hypothalamic regulation of body weight. *Endocr. Rev.* **20**, 68–100 (1999).
9. Cone, R. D. The central melanocortin system and energy homeostasis. *Trends Endocrinol. Metab.* **10**, 211–216 (1999).
10. Elias, C. F. *et al.* Leptin differentially regulates NPY and POMC neurons projecting to the lateral hypothalamic area. *Neuron* **23**, 775–786 (1999).
11. Glaum, S. R. *et al.* Leptin, the obese gene product, rapidly modulates synaptic transmission in the hypothalamus. *Mol. Pharmacol.* **50**, 230–235 (1996).
12. Spaniswick, D., Smith, M. A., Groppi, V. E., Logan, S. D. & Ashford, M. L. Leptin inhibits hypothalamic neurons by activation of ATP-sensitive potassium channels. *Nature* **390**, 521–525 (1997).
13. Lee, K., Dixon, A. K., Richardson, P. J. & Pinnock, R. D. Glucose-receptive neurones in the rat ventromedial hypothalamus express KATP channels composed of Kir6.1 and SUR1 subunits. *J. Physiol. (Lond.)* **515**, 439–452 (1999).
14. Shiraiishi, T., Sasaki, K., Nijima, A. & Oomura, Y. Leptin effects on feeding-related hypothalamic and peripheral neuronal activities in normal and obese rats. *Nutrition* **15**, 576–579 (1999).
15. Bagnol, D. *et al.* Anatomy of an endogenous antagonist: relationship between Agouti-related protein and proopiomelanocortin in brain. *J. Neurosci.* [online] (cited 25 Aug. 99) <<http://www.jneurosci.org/cgi/content/full/19/18/RC26>> (1999).
16. Butler, A. A. *et al.* A unique metabolic syndrome causes obesity in the melanocortin-3 receptor-deficient mouse. *Endocrinology* **141**, 3518–3521 (2000).
17. Young, J. I. *et al.* Authentic cell-specific and developmentally regulated expression of pro-opiomelanocortin genomic fragments in hypothalamic and hindbrain neurons of transgenic mice. *J. Neurosci.* **18**, 6631–6640 (1998).
18. Franklin, K. B. J. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* (Academic, San Diego, 1997).
19. Kelly, M. J., Loose, M. D. & Ronnekleiv, O. K. Opioids hyperpolarize β-endorphin neurons via μ-receptor activation of a potassium conductance. *Neuroendocrinology* **52**, 268–275 (1990).
20. Slugg, R. M., Hayward, M. D., Ronnekleiv, O. K., Low, M. J. & Kelly, M. J. Effect of the μ-opioid agonist DAMGO on medial basal hypothalamic neurons in β-endorphin knock-out mice. *Neuroendocrinology* **72**, 208–217 (2000).
21. Powis, J. E., Bains, J. S. & Ferguson, A. V. Leptin depolarizes rat hypothalamic paraventricular nucleus neurons. *Am. J. Physiol.* **274**, R1468–R1472 (1998).
22. Horvath, T. L., Bechmann, I., Naftolin, F., Kalra, S. P. & Leranth, C. Heterogeneity in the neuropeptide Y-containing neurons of the rat arcuate nucleus: GABAergic and non-GABAergic subpopulations. *Brain Res.* **756**, 283–286 (1997).
23. Broberger, C., Landry, M., Wong, H., Walsh, J. N. & Hokfelt, T. Subtypes Y1 and Y2 of the neuropeptide Y receptor are respectively expressed in pro-opiomelanocortin- and neuropeptide-Y-containing neurons of the rat hypothalamic arcuate nucleus. *Neuroendocrinology* **66**, 393–408 (1997).
24. King, P. J., Widdowson, P. S., Doods, H. N. & Williams, G. Regulation of neuropeptide Y release by neuropeptide Y receptor ligands and calcium channel antagonists in hypothalamic slices. *J. Neurochem.* **73**, 641–646 (1999).

25. Grieco, P., Balse, P. M., Weinberg, D., MacNeil, T. & Hruby, V. J. D-Amino acid scan of gamma-melanocyte-stimulating hormone: importance of Trp(8) on human MC3 receptor selectivity. *J. Med. Chem.* **43**, 4998–5002 (2000).
26. Cifary, A., Gorcs, T. J. & Palkovits, M. Neuropeptide Y innervation of ACTH-immunoreactive neurons in the arcuate nucleus of rats: a correlated light and electron microscopic double immunolabeling study. *Brain Res.* **506**, 215–222 (1990).
27. Horvath, T. L., Naftolin, F. & Leranth, C. GABAergic and catecholaminergic innervation of mediobasal hypothalamic beta-endorphin cells projecting to the medial preoptic area. *Neuroscience* **51**, 391–399 (1992).
28. Cowley, M. A. *et al.* Integration of NPY, AGRP, and melanocortin signals in the hypothalamic paraventricular nucleus: evidence of a cellular basis for the adipostat. *Neuron* **24**, 155–163 (1999).
29. Halaas, J. L. *et al.* Physiological response to long-term peripheral and central leptin infusion in lean and obese mice. *Proc. Natl Acad. Sci. USA* **94**, 8878–8883 (1997).

Acknowledgements

We wish to thank V. J. Hruby for the D-Trp⁸-γMSH, O. K. Ronnekleiv, R. G. Allen and M. R. Brown for antisera and J. T. Williams and J. M. Brundage for advice. This work was supported by the NIH, a Fogarty International Research Collaborative Award, the International Scholar Program of the Howard Hughes Medical Institute, and Agencia Nacional de Promoción Científica y Tecnológica.

Correspondence and requests for materials should be addressed to R.D.C. (e-mail: cone@ohsu.edu) or M.J.L. (e-mail: low@ohsu.edu).

Calmodulin bifurcates the local Ca²⁺ signal that modulates P/Q-type Ca²⁺ channels

Carla D. DeMaria*, Tuck Wah Soong†, Badr A. Alseikhan*, Rebecca S. Alvania* & David T. Yue*

* The Johns Hopkins University School of Medicine, Departments of Biomedical Engineering and Neuroscience, Program in Molecular and Cellular Systems Physiology, 720 Rutland Avenue, Baltimore, Maryland 21205, USA

† National Neuroscience Institute, 11 Jalan Tan Tuck Seng, Singapore 308433, and Department of Physiology, National University of Singapore

Acute modulation of P/Q-type (α_{1A}) calcium channels by neuronal activity-dependent changes in intracellular Ca²⁺ concentration may contribute to short-term synaptic plasticity^{1–3}, potentially enriching the neurocomputational capabilities of the brain^{4,5}. An unconventional mechanism for such channel modulation has been proposed^{6,7} in which calmodulin (CaM) may exert two opposing effects on individual channels, initially promoting ('facilitation') and then inhibiting ('inactivation') channel opening. Here we report that such dual regulation arises from surprising Ca²⁺-transduction capabilities of CaM. First, although facilitation and inactivation are two competing processes, both require Ca²⁺-CaM binding to a single 'IQ-like' domain on the carboxy tail of α_{1A}⁸; a previously identified 'CBD' CaM-binding site^{6,7} has no detectable role. Second, expression of a CaM mutant with impairment of all four of its Ca²⁺-binding sites (CaM₁₂₃₄) eliminates both forms of modulation. This result confirms that CaM is the Ca²⁺ sensor for channel regulation, and indicates that CaM may associate with the channel even before local Ca²⁺ concentration rises. Finally, the bifunctional capability of CaM arises from bifurcation of Ca²⁺ signalling by the lobes of CaM: Ca²⁺ binding to the amino-terminal lobe selectively initiates channel inactivation, whereas Ca²⁺ sensing by the carboxy-terminal lobe induces facilitation. Such lobe-specific detection provides a compact means to decode local Ca²⁺ signals in two ways, and to separately initiate distinct actions on a single molecular complex.

To simplify the dissection of the molecular mechanisms, we studied recombinant P/Q-type (α_{1A}/β_{2A}/α_{2δ}) channels expressed in mammalian HEK293 cells. Figure 1a shows that Ca²⁺-dependent

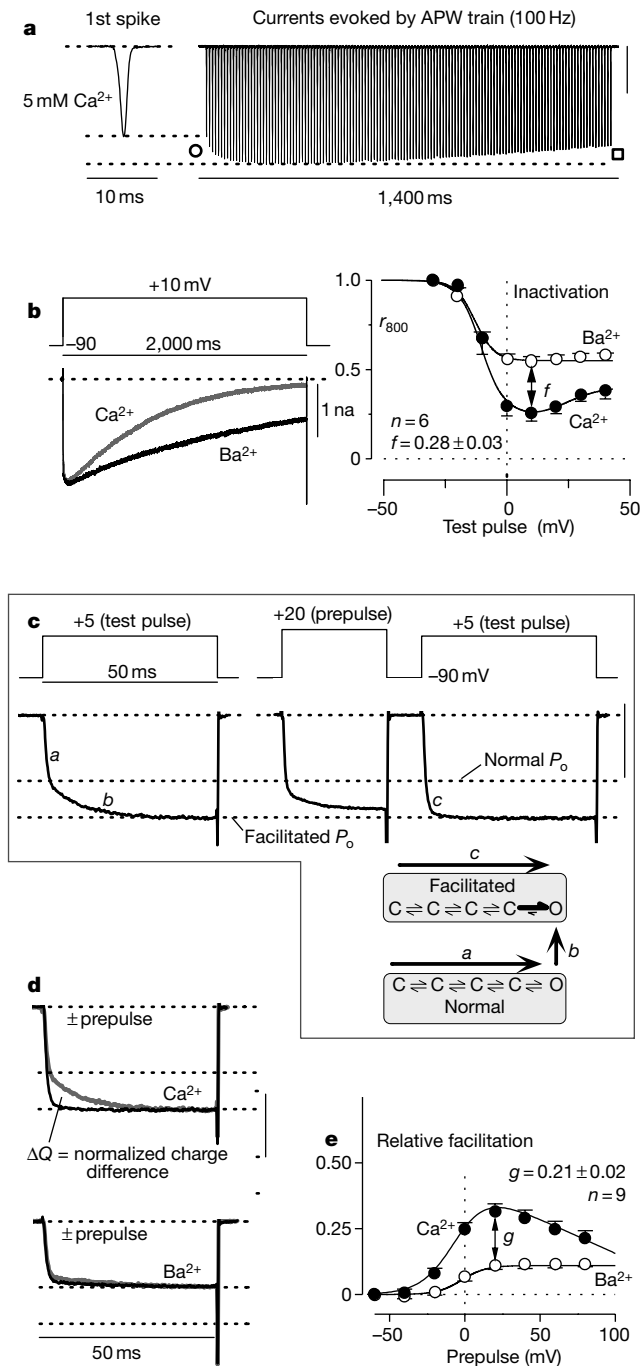


Figure 1 Facilitation and inactivation in P/Q-type Ca^{2+} channels. **a**, Ca^{2+} currents show facilitation (circle) and inactivation (square) during trains of action potential waveforms (APW). **b**, Left, inactivation of Ca^{2+} (grey) and Ba^{2+} (black) currents during step depolarization. Ca^{2+} traces for inactivation amplified about twice to match Ba^{2+} traces, and tail currents clipped at borders. Ba^{2+} scale, 1 nA. Right, r_{800} , averaged from n cells. f was taken at 10 mV. **c**, Ca^{2+} -dependent facilitation in prepulse protocols. Left, test pulse without prepulse. Right, test pulse following prepulse, with rapid activation to larger facilitated level. Ca^{2+} scale, 1 nA. P_o , steady-state open probability. Same cell as in **b** was used. Bottom, interpretation of phases *a*–*c*. *a*, Initial rapid activation; *b*, slower phase of increase; *c*, the channels already in the facilitated mode rapidly activate. **d**, Quantification of facilitation. Top, test-pulse Ca^{2+} currents obtained without (grey) and with (black) prepulse, from **c** after normalization to unity at test-pulse termination. ΔQ , integral of difference between normalized traces. Relative facilitation $\text{RF} = \Delta Q/\tau = 0.33$ for upper traces, where τ is the facilitation time constant (~ 10 ms) (see Methods). Bottom, test-pulse Ba^{2+} currents, same cell ($\text{RF} = 0.11$). Scale, 2 nA. Ba^{2+} traces scaled to reflect lower open probability of normal mode. **e**, Average relative facilitation. g taken at 20 mV. **a**–**e**, See Supplementary Information.

facilitation and inactivation in such recombinant channels recapitulate modulatory behaviour compatible with that of presynaptic channels^{1,2}. To mimic physiological responses, we activated recombinant channels using trains of action-potential waveforms. The resulting Ca^{2+} currents facilitated with repetitive spikes, and then inactivated over a longer timescale. Such behaviour was absent in corresponding Ba^{2+} currents (see Supplementary Information), fitting with the high selectivity of CaM^8 for Ca^{2+} over Ba^{2+} .

To quantify the inactivation, we used prolonged square-pulse depolarizations (Fig. 1b, left), in which Ca^{2+} current decayed almost completely whereas the corresponding Ba^{2+} current remained substantial. On average (Fig. 1b, right), the fraction of peak Ca^{2+} current present after depolarizing for 800 ms (r_{800}) bore a deep, U-shaped dependence upon test-pulse voltage, consistent with genuine Ca^{2+} -dependent inactivation⁸. The corresponding Ba^{2+} relation declined only modestly, reflecting a slower voltage-dependent mechanism. Hence, the difference between Ca^{2+} and Ba^{2+} $r_{800}(f)$ provided a robust index of pure Ca^{2+} -dependent inactivation.

Using shorter pulses with negligible inactivation (Fig. 1c, left top), facilitation was readily resolved as a slower phase of Ca^{2+} current increase (*b*), following an initial rapid activation (*a*). The phenomenon could be understood as fast activation in a normal mode of gating (Fig. 1c, bottom, pathway *a*), followed by slower, Ca^{2+} -driven conversion to a facilitated gating mode with enhanced open probability (pathway *b*). Such a scenario was confirmed in three ways. First, a voltage prepulse should ‘prefacilitate’ the

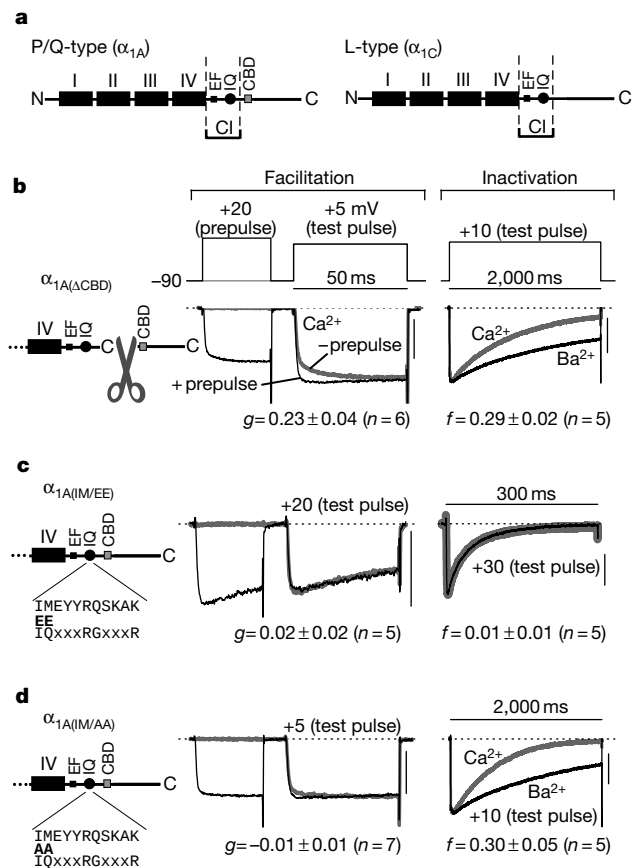


Figure 2 Structural determinants of facilitation and inactivation. **a**, P/Q-type and L-type channel pore-forming α_1 -subunits. Channel domains (I–IV). Proximal third of C-tail (CI region) contains consensus EF-hand (EF), IQ motif (IQ) and CBD domain. **b**–**d**, Format essentially as in Fig. 1, except **c**. Scales, 1 nA, except prepulse in **c** (0.5 nA). **c**, IQ consensus⁸ shown for reference. Higher charge carrier (Methods) necessitated different test-pulse voltages. Faster inactivation required shorter inactivation test pulses, and use of r_{100} (fraction of peak current after a 100-ms depolarization) to quantify inactivation. f , difference in Ba^{2+} and Ca^{2+} r_{100} at 30 mV. **a**–**d**, See Supplementary Information.

channels, such that a subsequent test pulse would rapidly activate the channels already in the facilitated gating mode (Fig. 1c, right and bottom, pathway c). The overlay of normalized test-pulse Ca^{2+} currents (obtained $\pm$ prepulse) underscores such prefacilitation (Fig. 1d, top). Second, test-pulse Ba^{2+} currents should activate by a single, rapid component regardless of prepulse depolarization (Fig. 1d, bottom), because of channel 'trapping' in the normal mode. Finally, if Ca^{2+} entry drives facilitation, then prefacilitation should bear a bell-shaped dependence upon prepulse potential (Fig. 1e). This was confirmed by the average relation between relative facilitation and prepulse voltage. Here, the relative facilitation is proportional to the fraction of channels facilitated by a prepulse, as relative facilitation is derived from the difference in charge (ΔQ ; Fig. 1d) carried by normalized test-pulse currents (elicited $\pm$ prepulse; Fig. 1d). The corresponding Ba^{2+} relative facilitation (Fig. 1e) showed only a small monotonic increase, reflecting background G-protein modulation¹⁰. Hence, the difference between Ca^{2+} and Ba^{2+} relative facilitations (g; Fig. 1e) gave a measure of pure Ca^{2+} -dependent facilitation.

How does CaM initiate such kinetically disparate processes—facilitation and inactivation—on a single target molecule? One possibility is that Ca^{2+} -CaM binds to two channel sites, one for each modulatory effect. Although the CaM-binding domain (CBD) site on the distal carboxyl tail (C-tail) of α_{1A} reportedly mediated both facilitation and inactivation^{6,7} (Fig. 2a, left), CaM also binds to an upstream IQ-like motif⁸. Though the function of the α_{1A} IQ site was unknown, CaM binding to the homologous L-type channel IQ site (Fig. 2a, right) is crucial to Ca^{2+} regulation^{8,11–13}, and the encompassing CI (calcium inactivation) region may transduce Ca^{2+} -CaM binding to L-type channel modulation¹⁴.

To test for a functional role of sites outside the CBD, we truncated the α_{1A} C-tail so as to entirely remove the CBD region ($\alpha_{1A}(\Delta\text{CBD})$; Fig. 2b). Unexpectedly, facilitation and inactivation were not appreciably changed compared with the wild-type α_{1A} , implicating other CaM-binding sites in channel regulation. To test the importance of the α_{1A} IQ site, we mutated the critical isoleucine and methionine residues of the IQ-like motif¹⁵ to negatively charged glutamates ($\alpha_{1A}(\text{IM/EE})$; Fig. 2c), in an attempt to disrupt Ca^{2+} -CaM interaction. These mutations entirely abolished facilitation and inactivation, suggesting that both regulatory effects are initiated through Ca^{2+} -CaM interaction with the IQ site. The residual voltage-dependent inactivation, identical with either Ca^{2+} or Ba^{2+} , was accelerated about 40-fold. This fit with the dual roles of transduction and CaM binding served by many CaM-interaction domains¹⁶ (see Supplementary Information). We also examined more modest mutations to neutrally charged alanines ($\alpha_{1A}(\text{IM/AA})$; Fig. 2d), in an attempt to preserve partial Ca^{2+} -CaM interaction. Such mutations eliminated facilitation, but spared Ca^{2+} -dependent inactivation.

The presumed effects of channel mutations on Ca^{2+} -CaM interaction were verified by gel-mobility shift assays, in which mixtures of CaM and peptides spanning the α_{1A} IQ region were electrophoresed on nondenaturing gels (Fig. 3a–c). Without Ca^{2+} , CaM ran as a single, peptide-free band, regardless of the concentration of the wild-type peptide IQ_{WT} (Fig. 3a, left). By contrast, with Ca^{2+} , increasing peptide concentrations induced a graded interchange towards a slower mobility band (Fig. 3a, right), representing a peptide- Ca^{2+} -CaM complex. As we observed previously⁸, the mobility shift saturated at low peptide:CaM ratios, indicating a substantial Ca^{2+} -CaM affinity for the peptide. By contrast, the $\text{IQ}_{\text{IM/EE}}$ peptide (Fig. 3b) showed no indication of peptide-CaM interaction, whereas the $\text{IQ}_{\text{IM/AA}}$ peptide (Fig. 3c) showed a partial mobility shift, indicating a modest interaction with Ca^{2+} -CaM. Fluorescence experiments with dansyl-CaM¹⁷ (Fig. 3d) quantitatively confirmed this rank order of peptide- Ca^{2+} -CaM interaction. In saturating Ca^{2+} , IQ_{WT} bound CaM with high affinity ($K_d = 66 \text{ nM}$), $\text{IQ}_{\text{IM/AA}}$ bound with intermediate affinity

($K_d = 1.03 \mu\text{M}$) and $\text{IQ}_{\text{IM/EE}}$ showed no appreciable interaction. These biochemical studies (Fig. 3) confirmed the presumptions of functional mutational analysis in Fig. 2 (that IM/EE would eliminate, and IM/AA would weaken Ca^{2+} -CaM interaction), leading us to argue that Ca^{2+} -CaM binding to the α_{1A} IQ site alone mediates both facilitation and inactivation. The initial question about how the interaction of CaM with a single CBD site produces two modulatory effects was then replaced by a similar puzzle involving the IQ site.

The approach for solving this paradox came by first considering the fast kinetics of facilitation ($\tau \approx 10 \text{ ms}$, Fig. 1c). Such rapidity seems incompatible with the diffuse cytoplasmic localization typical of CaM. Might apoCaM (Ca^{2+} -free CaM) be bound to the channel complex even before Ca^{2+} entry ('preassociation'), thus accelerating

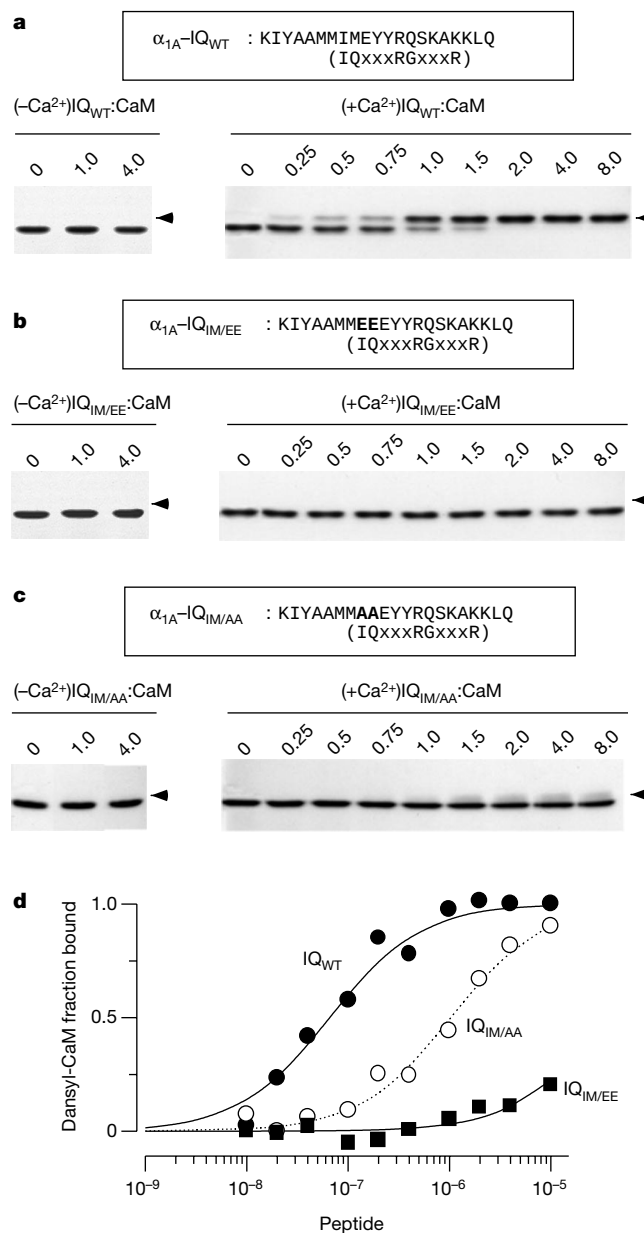


Figure 3 CaM binding to α_{1A} IQ site. **a–c**, Gel-mobility shift assay of CaM binding to peptides (insets) spanning the α_{1A} IQ site. Without Ca^{2+} (**a–c**, left), lack of CaM mobility shift. With Ca^{2+} (**a–c**, right), increasing IQ_{WT} :CaM ratios (**a**) caused exchange towards slower-mobility (peptide- Ca^{2+} -CaM) band (triangle); $\text{IQ}_{\text{IM/EE}}$ induced no shift in Ca^{2+} (**b**), and $\text{IQ}_{\text{IM/AA}}$ produced partial shift in Ca^{2+} (**c**). **d**, Dansyl-CaM fluorescence characterization of Ca^{2+} -CaM binding to peptides (see Methods). Smooth curves, single-binding isotherm fits.

subsequent Ca^{2+} -CaM interaction with the IQ site? To examine this question we coexpressed P/Q-type channels with CaM_{1234} (Fig. 4a, top), a Ca^{2+} -insensitive CaM mutant with aspartate-to-alanine mutations at the 'x' position of all four EF hands^{8,18}. CaM_{1234} would eliminate channel regulation if the CaM responsible for modulation was prebound to the α_{1A} complex with preferential access to the IQ site; otherwise, functional endogenous CaM should preserve normal channel regulation⁸. The observed elimination of facilitation and inactivation (Fig. 4a) indicates preassociation of apoCaM, and unequivocally confirms that CaM is the Ca^{2+} sensor for regulation.

The dominant-negative action of CaM_{1234} allowed us to return to the question of how Ca^{2+} -CaM binding to a single IQ site initiates dual modulatory processes. Mutant CaMs with selective impairment of Ca^{2+} binding to either the N-terminal (CaM_{12}) or C-terminal lobe (CaM_{34}) (Fig. 4b, c, left) provided the key insight. Expression of CaM_{12} , in which the aspartate-to-alanine mutations described for CaM_{1234} were restricted to the N-terminal lobe, strikingly suppressed Ca^{2+} -dependent inactivation (Fig. 4b), whereas facilitation was indistinguishable from the control (Fig. 1e). Expression of CaM_{34} (Fig. 4c), in which the analogous mutations

were restricted to the C-terminal lobe, yielded the opposite results: Ca^{2+} -dependent facilitation was essentially absent, whereas inactivation remained unchanged compared with the control (Fig. 1b). These unexpected findings were confirmed by additional population data (Fig. 4d) that gave the indices for the strength of inactivation (f) and facilitation (g); these were pooled from even more cells than were used for the full voltage protocols. We therefore conclude that Ca^{2+} binding to the N-terminal lobe of CaM selectively initiates inactivation, whereas Ca^{2+} binding to the C-terminal lobe selectively triggers facilitation.

Our experiments highlight signalling capabilities of CaM that may generalize to numerous biological systems. First, CaM regulation through the IQ-like site of α_{1A} adds to an emerging theme, as a homologous IQ site in L-type (α_{1C}) Ca^{2+} channels underlies their regulation by CaM^{8,11-14}, and analogous IQ-like motifs in R-type (α_{1E}) Ca^{2+} channels⁸ and Na channels¹⁹ interact with CaM. Second, preassociation of apoCaM with molecular complexes may be a prevalent mechanism to speed their modulation by Ca^{2+} -CaM. Both L-type Ca^{2+} channels^{8,11} and small-conductance K channels^{20,21} may exploit such a scheme. Moreover, FRET (fluorescence resonance energy transfer) experiments in our laboratory indicate a

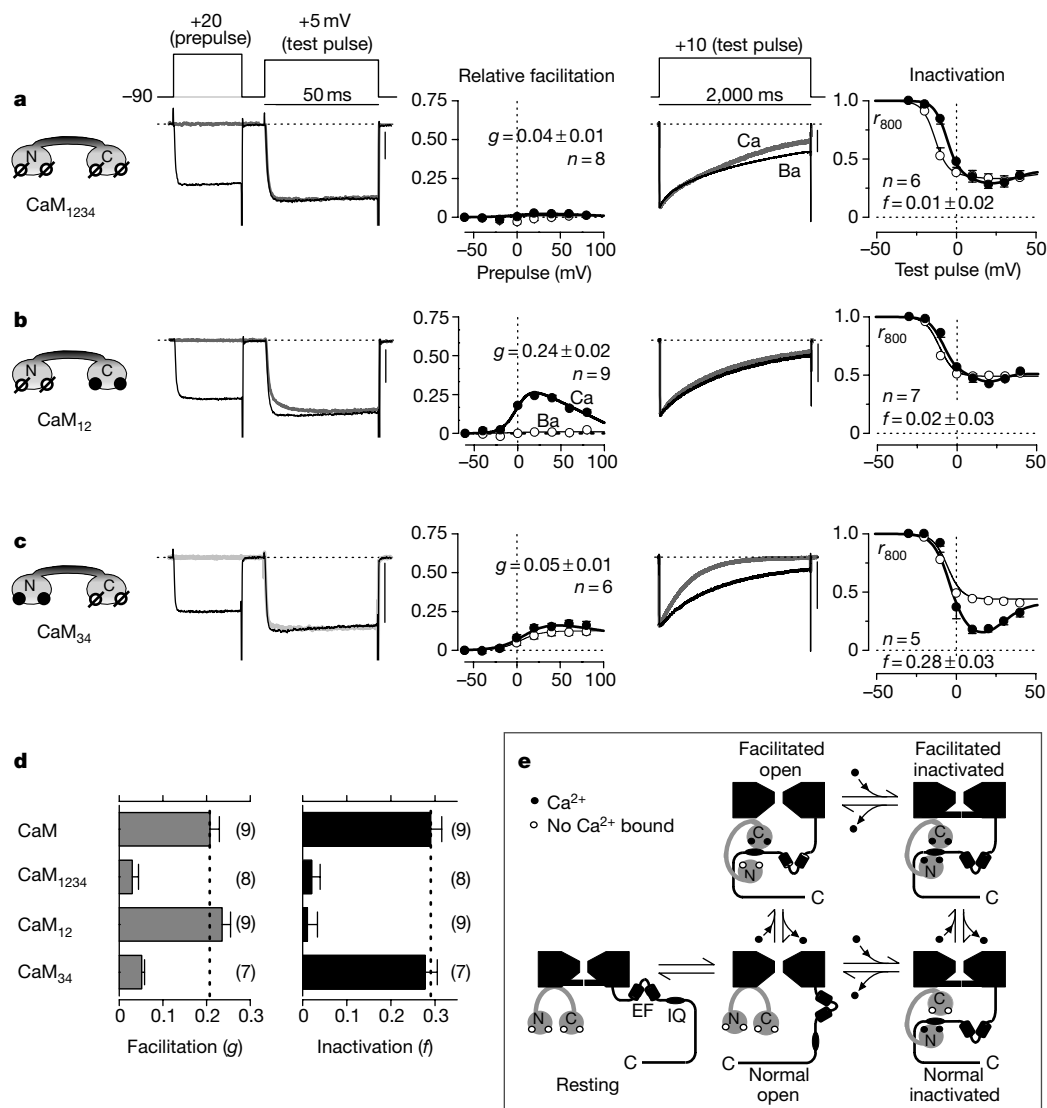


Figure 4 Selective initiation of facilitation and inactivation by different lobes of CaM. **a–c**, Format as in Fig. 1. Prepulse and Ba^{2+} inactivation scale, 1 nA. **a**, Co-expression of CaM_{1234} with channels. **b**, Co-expression of CaM_{12} (defective N-terminal lobe) with channels. **c**, Co-expression of CaM_{34} (defective C-terminal lobe) with channels.

d, Summarized effects of various CaMs. CaM indicates endogenous CaM without recombinant expression. For each CaM, the same set of cells was used for facilitation and inactivation. **e**, Proposed mechanism of P/Q-type channel regulation. **a–e**, See Supplementary Information.

preassociation of CaM to multiple types of Ca^{2+} channels in living cells²², and apoCaM binding sites have been identified in Na channels¹⁹ and ryanodine receptors²³. Third, our results show that CaM, acting through a single binding region, can use lobe-specific Ca^{2+} binding to selectively control distinct processes, perhaps as proposed for P/Q-type channels (Fig. 4e). ApoCaM preassociates with resting channels at a saturable site that provides preferential access to the IQ region. Depolarization opens channels, and the ensuing Ca^{2+} binding to the C-terminal lobe of CaM enables one form of interaction with the IQ site, which induces facilitated channel conformations. Ca^{2+} binding to the N-terminal lobe enables a second form of interaction that favours inactivated conformations. Selective elimination of inactivation or facilitation by CaM_{12} and CaM_{34} (Fig. 4d) shows that the processes are distinct. Such a mechanism could provide intriguing design advantages, especially because it seems likely that the lobes of CaM selectively detect features of the local Ca^{2+} signal that arise from local and distant sources of Ca^{2+} (Fig. 5). The C-terminal lobe may respond preferentially to the spike-like component of local Ca^{2+} concentration that is attributable to local channel activity (C-lobe CaM readout). The N-terminal lobe may detect the slow component of the signals which result from aggregate cellular Ca^{2+} signalling (N-lobe CaM readout) (see Supplementary Information). There are suggestive biochemical precedents for lobe-specific CaM signalling^{16,24}, and functional precedents where one lobe of CaM preferentially regulates *Paramecium*²⁵ and mammalian^{8,20} channels, as well as yeast physiology²⁶. However, the P/Q-type channel may provide the first example of CaM lobe-specific initiation of two regulatory effects on a single molecular target complex. If widespread, such bifurcation of local Ca^{2+} signals by CaM may be a critical rationale for the two-lobed design of calmodulin. □

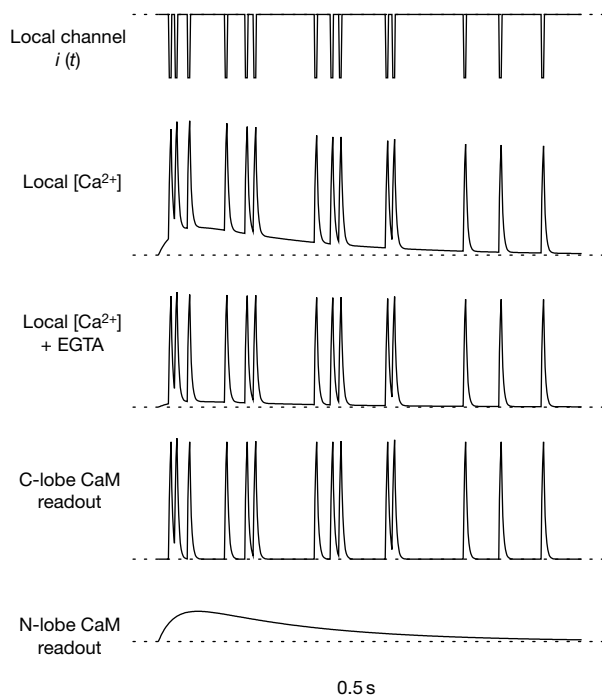


Figure 5 Decoding local Ca^{2+} concentrations ($[\text{Ca}^{2+}]$). Dotted lines, zero. Local channel $i(t)$, schematized currents of one Ca^{2+} channel. Theoretical local $[\text{Ca}^{2+}]$ near channel³⁰, reflecting activity of associated channel (spikes) and distant Ca^{2+} sources (slow phase). Local $[\text{Ca}^{2+}] + \text{EGTA}$, postulated local Ca^{2+} with several mM cytoplasmic EGTA, featuring attenuation of slow phase with sparing of spikes³⁰. C-lobe CaM readout, spike component. Deduction from sparing of facilitation with mM EGTA⁷, facilitation dependence on CaM C-terminal lobe, and third trace. N-lobe CaM readout, slow phase. Deduction from inactivation blunting by mM EGTA⁷, inactivation dependence on CaM N-terminal lobe, and third trace. Approximate timebase. See Supplementary Information.

Methods

Molecular biology

Mutations of the α_{1A} IQ-like region⁸ (Fig. 2a, IMEYYRQSKAK) were made by PCR, with human α_{1A} as template²⁷. For $\alpha_{1A}(\text{IM/EE})$ and $\alpha_{1A}(\text{IM/AA})$ mutants (Fig. 2c, d), forward mutagenic oligonucleotides spanned both the IQ region and a *Bgl*III site, 20 bp upstream of the reference isoleucine (bold). The reverse oligo had an *Xba*I site and sequence complementary to the α_{1A} stop codon. The resulting 899-base-pair (bp) PCR products were subcloned into human $\alpha_{1A}/\text{pcDNA3}$ (Invitrogen) by *Bgl*III–*Xba*I. For $\alpha_{1A}(\Delta\text{ICBD})$ (Fig. 2b), the forward oligo annealed 44 bp upstream of an *Xho*I site preceding IVS6. The reverse oligo was complementary to the amino-acid stretch LDP, 51 amino acids downstream from the reference isoleucine. This was followed by a premature stop codon and *Xba*I site. The resulting 779-bp PCR fragment was subcloned into human $\alpha_{1A}/\text{pcDNA3}$ by *Xho*I–*Xba*I. All PCR products were entirely sequenced. CaMs were cloned into pcDNA3 (ref. 8).

Gel-mobility-shift assay

CaM gel-shift assays were performed with 15% non-denaturing polyacrylamide gels mostly as described⁸. In Ca^{2+} -free experiments, 5 mM EGTA was added to reaction buffers and 2 mM EGTA was included in the running buffer, but no chelator or exogenous Ca^{2+} was added during gel casting. In experiments with Ca^{2+} , 1–2 mM Ca^{2+} was present in the reaction buffer and 0.1 mM Ca^{2+} was included in both the running buffer and gel (during casting).

Dansyl-CaM studies

Purified recombinant CaM protein was derivatized²⁸ with dansyl (5-dimethylaminonaphthalene-1-sulphonyl) chloride (Molecular Probes). The dansyl-CaM was then dialysed against 20 mM MOPS (pH 7.2). During binding assays, 100 nM dansyl-CaM was mixed with various concentrations of α_{1A} -IQ_{WT} peptide (Fig. 3a), in a room-temperature buffer¹⁷ containing 2 mM CaCl_2 , 150 mM NaCl and 50 mM Tris-Cl (pH 7.5). Comparable results were obtained with 150 μM CaCl_2 (not shown). Fluorescence emission (400–650 nm, 20 nm bandwidth) was monitored by a spectrofluorometer (SPF-500C, SLM Instruments), using 340 nm excitation (2.5 nm bandwidth). Background-subtracted fluorescence emission at 490 nm (F_{490}) specified the apparent fraction of CaM bound to peptide (B_{app}), with the relation $B_{\text{app}} = (F_{490} - F_{490}[\text{no peptide}]) / (F_{490}[\text{10 } \mu\text{M peptide}] - F_{490}[\text{no peptide}])$. This algorithm exploits a peptide-dependent blue shift and enhancement of dansyl emission spectra¹⁷. The relations between B_{app} and $[\alpha_{1A}\text{-IQ}_{\text{WT}}$ peptide] were least-squares fit with the relation $B_{\text{app}} = B_{\text{app,max}} / (K_d + [\alpha_{1A}\text{-IQ}_{\text{WT}}$ peptide]), and the actual fraction of CaM bound to peptide determined as $B = B_{\text{app}} / B_{\text{app,max}}$ (Fig. 3d). We applied an analogous process to the relation between B_{app} and $[\alpha_{1A}\text{-IQ}_{\text{IM/AA}}$ peptide]. For the $\alpha_{1A}\text{-IQ}_{\text{IM/EE}}$ peptide, the equation for B_{app} used a 100- μM peptide measurement (instead of a 10- μM). The peptides were from HHMI Biopolymer Laboratory, Johns Hopkins University School of Medicine.

Electrophysiology

The methods here were mostly as described⁸. cDNA for wild-type (or mutant) human α_{1A} (ref. 27) was transiently cotransfected with β_{2A} and α_{2B} (ref. 8) (and various CaMs⁸ as required) in HEK293 cells. Two to three days later, whole-cell recordings were obtained at room temperature. β_{2A} minimized voltage inactivation²⁹, enhancing resolution of Ca^{2+} -dependent regulation. Bath solution contained (in mM): TEA- MeSO_3 , 140; HEPES (pH 7.3), 10; and CaCl_2 or BaCl_2 , 5; at 300 mOsm, adjusted with glucose. Internal solution (in mM): Cs- MeSO_3 , 135; CsCl₂, 5; EGTA, 0.5; MgCl₂, 1; Mg-ATP, 4; and HEPES (pH 7.3), 10; at 290 mOsm, adjusted with glucose. To enhance the resolution of $\alpha_{1A}(\text{IM/EE})$ (Fig. 2c), bath solution contained 20 mM BaCl_2 or CaCl_2 and 119 mM TEA- MeSO_3 . Larger surface-potential shift caused by 20 mM charge carrier required increased test-pulse voltages (Fig. 2c). The protocols were otherwise the same as for the experiments with 5 mM charge carrier. For action potential waveforms (APWs) (Fig. 1a), voltages were precorrected for a –11 mV junction potential¹⁰, with seals made in Ba^{2+} bath solution; otherwise, reported voltages were uncorrected, and true voltage may be obtained by subtracting 11 mV from reported values. APWs scaled uniformly from those recorded in calyx of Held²⁹, here with a 2-ms half width and a voltage range from –80 to 34 mV. Currents were filtered at 2 kHz and sampled at 10 kHz, except in APW experiments (5 kHz lowpass, 25 kHz sampling). Series resistance was typically 1–2 M Ω after more than 70% compensation. Leaks and capacitive transients were subtracted by P/8 protocol. Test-pulse depolarizations were delivered every 60 s (facilitation protocol) or 100 s (inactivation protocol).

Ca^{2+} -dependent facilitation was determined using the normalized charge difference ΔQ , obtained by integrating the difference between normalized traces $\pm$ prepulse (Fig. 1d, top). The fraction of channels facilitated by prepulse ($F_{\text{facilitated}}$) is directly proportional to ΔQ divided by the time constant (τ) of facilitation, yielding relative facilitation ($\text{RF} = \Delta Q/\tau$). This follows by assuming that all channels are initially in normal mode at test-pulse onset, and that subsequent shifts to facilitated mode occur mono-exponentially with a time constant τ . Then, $\text{RF} = F_{\text{facilitated}} \times [P_{\text{o,facilitated}} - P_{\text{o,normal}}] / P_{\text{facilitated}}$, where $P_{\text{o,facilitated}}$ and $P_{\text{o,normal}}$ are steady-state open probabilities in facilitated and normal modes, respectively. τ was explicitly determined from Ca^{2+} traces in each cell before calculation of RF . Ba^{2+} RF calculated by using τ values determined from Ca^{2+} traces in the same cell. For knockouts of facilitation (Figs 2c, d, 4a, c), τ was set to 10 ms (at about the average for control facilitation, Fig. 1d) in RF calculations.

All average data were presented as mean $\pm$ s.e.m., after analysis by custom-written software in MATLAB (MathWorks). Smooth-curve fits to data were done by eye, except in Fig. 3d. In Fig. 4, an important feature was to include only cells in which both g and f indices for facilitation and inactivation were determined. This constraint excluded the

possibility that selection bias might have led to an apparent differential elimination of regulatory mechanisms. Such an artefact could have occurred if mutant CaMs variably inhibited facilitation, inactivation, or both—depending on the particular cell under observation. Appropriate changes (or lack thereof) in f and g values, as determined from the same cells, excluded such a scenario (Fig. 4; see Supplementary Information).

Received 14 December 2000; accepted 30 March 2001.

- Borst, J. G. & Sakmann, B. Facilitation of presynaptic calcium currents in the rat brainstem. *J. Physiol. (Lond.)* **513**, 149–155 (1998).
- Cuttle, M. F., Tsujimoto, T., Forsythe, I. D. & Takahashi, T. Facilitation of the presynaptic calcium current at an auditory synapse in rat brainstem. *J. Physiol. (Lond.)* **512**, 723–729 (1998).
- Forsythe, I. D., Tsujimoto, T., Barnes-Davies, M., Cuttle, M. & Takahashi, T. Inactivation of presynaptic calcium current contributes to synaptic depression at a fast central synapse. *Neuron* **20**, 797–807 (1998).
- Abbott, L. F., Varela, J. A., Sen, K. & Nelson, S. B. Synaptic depression and cortical gain control. *Science* **175**, 220–224 (1997).
- Tsodyks, M. V. & Markram, H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl Acad. Sci. USA* **94**, 719–723 (1997).
- Lee, A. *et al.* Ca^{2+} /calmodulin binds to and modulates P/Q-type calcium channels. *Nature* **399**, 155–159 (1999).
- Lee, A., Scheuer, T. & Catterall, W. A. Ca^{2+} /calmodulin-dependent facilitation and inactivation of P/Q-type Ca^{2+} channels. *J. Neurosci.* **20**, 6830–6838 (2000).
- Peterson, B. Z., DeMaria, C. D., Adelman, J. P. & Yue, D. T. Calmodulin is the Ca^{2+} sensor for Ca^{2+} -dependent inactivation of L-type calcium channels. *Neuron* **22**, 549–558 (1999).
- Chao, S. H., Suzuki, Y., Zysk, J. R. & Cheung, W. Y. Activation of calmodulin by various metal cations as a function of ionic radius. *Mol. Pharmacol.* **26**, 75–82 (1984).
- Colecraft, H. M., Patil, P. G. & Yue, D. T. Differential occurrence of reluctant openings in G-protein-inhibited N- and P/Q-type calcium channels. *J. Gen. Physiol.* **115**, 175–192 (2000).
- Zuhlke, R. D., Pitt, G. S., Deisseroth, K., Tsien, R. W. & Reuter, H. Calmodulin supports both inactivation and facilitation of L-type calcium channels. *Nature* **399**, 159–162 (1999).
- Zuhlke, R. D., Pitt, G. S., Tsien, R. W. & Reuter, H. Ca^{2+} -sensitive inactivation and facilitation of L-type Ca^{2+} channels both depend on specific amino acid residues in a consensus calmodulin-binding motif in the α_{1C} subunit. *J. Biol. Chem.* **275**, 21121–21129 (2000).
- Qin, N., Olcese, R., Bransby, M., Lin, T. & Birnbaumer, L. Ca^{2+} -induced inhibition of the cardiac Ca^{2+} channel depends on calmodulin. *Proc. Natl Acad. Sci. USA* **96**, 2435–2438 (1999).
- Peterson, B. Z. *et al.* Critical determinants of Ca^{2+} -dependent inactivation within an EF-hand motif of L-type Ca^{2+} channels. *Biophys. J.* **78**, 1906–1920 (2000).
- Houdusse, A. & Cohen, C. Target sequence recognition by the calmodulin superfamily: implications from light chain binding to the regulatory domain of scallop myosin. *Proc. Natl Acad. Sci. USA* **92**, 10644–10647 (1995).
- Elshorst, B. *et al.* NMR solution structure of a complex of calmodulin with a binding peptide of the Ca^{2+} pump. *Biochemistry* **38**, 12320–12332 (1999).
- Ehlers, M. D., Zhang, S., Bernhardt, J. P. & Huganir, R. L. Inactivation of NMDA receptors by direct interaction of calmodulin with the NR1 subunit. *Cell* **84**, 745–755 (1996).
- Putkey, J. A., Sweeney, H. L. & Campbell, S. T. Site-directed mutation of the trigger calcium-binding sites in cardiac troponin C. *J. Biol. Chem.* **264**, 12370–12378 (1989).
- Mori, M. *et al.* Novel interaction of the voltage-dependent sodium channel (VDSC) with calmodulin: does VDSC acquire calmodulin-mediated Ca^{2+} -sensitivity? *Biochemistry* **39**, 1316–1323 (2000).
- Keen, J. E. *et al.* Domains responsible for constitutive and Ca^{2+} -dependent interactions between calmodulin and small conductance Ca^{2+} -activated potassium channels. *J. Neurosci.* **19**, 8830–8838 (1999).
- Fanger, C. M. *et al.* Calmodulin mediates calcium-dependent activation of the intermediate conductance KCa channel, IKCa1. *J. Biol. Chem.* **274**, 5746–5754 (1999).
- Erickson, M. G. & Yue, D. T. FRET reveals tethering of calmodulin to calcium channel complex in single living cells. *Biophys. J.* **80**, 196a (2001).
- Rodney, G. G. *et al.* Calcium binding to calmodulin leads to an N-terminal shift in its binding site on the ryanodine receptor. *J. Biol. Chem.* **276**, 2069–2074 (2001).
- Barth, A., Martin, S. R. & Bayley, P. M. Specificity and symmetry in the interaction of calmodulin domains with the skeletal muscle myosin light chain kinase target sequence. *J. Biol. Chem.* **273**, 2174–2183 (1998).
- Kink, J. A. *et al.* Mutations in paramyosin calmodulin indicate functional differences between the C-terminal and N-terminal lobes in vivo. *Cell* **62**, 165–174 (1990).
- Ohya, Y. & Botstein, D. Diverse essential functions revealed by complementing yeast calmodulin mutants. *Science* **263**, 963–966 (1994).
- Sutton, K. G., McRory, J. E., Guthrie, H., Murphy, T. H. & Snutch, T. P. P/Q-type calcium channels mediate the activity dependent feedback of syntaxin-1A. *Nature* **401**, 800–804 (1999).
- Kincaid, R. L., Billingsley, M. L. & Vaughan, M. Preparation of fluorescent, cross-linking, and biotinylated calmodulin derivatives and their use in studies of calmodulin-activated phosphodiesterase and protein phosphatase. *Methods Enzymol.* **159**, 605–626 (1988).
- Patil, P. G., Brody, D. L. & Yue, D. T. Preferential closed-state inactivation of neuronal calcium channels. *Neuron* **20**, 1027–1038 (1998).
- Song, L. S., Sham, J. S., Stern, M. D., Lakatta, E. G. & Chang, H. Direct measurement of SR release flux by tracking Ca^{2+} spikes in rat cardiac myocytes. *J. Physiol. (Lond.)* **512**, 677–691 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank W. Agnew, C. Chen, H. Colecraft, M. Erickson, S. Takahashi, H. Agler and E. Sobie for discussion; B. Peterson for initial attempts to detect P/Q-type channel facilitation; and T. Snutch for the gift of human α_{1A} clone. This work was supported by an NIH NRSA fellowship (C.D.D.) and grants from the NIH (D.T.Y.) and NNI (T.W.S.).

Correspondence and requests for materials should be addressed to D.T.Y. (e-mail: dyue@bme.jhu.edu).

B cells acquire antigen from target cells after synapse formation

Facundo D. Batista, Dagmar Iber & Michael S. Neuberger

Medical Research Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, UK

Soluble antigen binds to the B-cell antigen receptor and is internalized for subsequent processing and the presentation of antigen-derived peptides to T cells¹. Many antigens are not soluble, however, but are integral components of membrane; furthermore, soluble antigens will usually be encountered *in vivo* in a membrane-anchored form, tethered by Fc or complement receptors^{2–4}. Here we show that B-cell interaction with antigens that are immobilized on the surface of a target cell leads to the formation of a synapse and the acquisition, even, of membrane-integral antigens from the target. B-cell antigen receptor accumulates at the synapse, segregated from the CD45 co-receptor which is excluded from the synapse, and there is a corresponding polarization of cytoplasmic effectors in the B cell. B-cell antigen receptor mediates the gathering of antigen into the synapse and its subsequent acquisition, thereby potentiating antigen processing and presentation to T cells with high efficacy. Synapse formation and antigen acquisition will probably enhance the activation of B cells at low antigen concentration, allow context-dependent antigen recognition and enhance the linking of B- and T-cell epitopes.

To investigate the B-cell response to antigen encountered as part of an immune complex tethered to a cell surface, immune complexes comprising hen-egg lysozyme (HEL) aggregated with specific immunoglobulin- γ (IgG) monoclonal antibodies were loaded on to the surface of an Fc γ receptor (Fc γ R)-expressing myeloid cell line. The immune complexes had a patchy distribution over the myeloid cell surface (Fig. 1a, b); however, on incubation with antigen-specific B cells, cell aggregates were formed in which the immune complexes on the myeloid cell and the B-cell antigen receptor (BCR) on the B cell were gathered together into a region of synapsis (Fig. 1c, d).

Similar results were obtained using immune complexes loaded onto Fc γ RI-expressing L-cell transfectants. Immunocytochemistry suggested that there is a gathering of tethered immune complexes, mediated by the BCR (possibly aided by the oligomeric nature of the BCR⁵) and accompanied by an apparent reorganization of the B-cell surface, as judged by segregation of the BCR from CD45 in the region of synapsis (Fig. 1e).

This concentration of BCR is reminiscent of the capping of surface IgM that results from incubation of B cells with polyvalent anti-IgM antisera^{6,7}. We therefore determined whether the reorganization of the B cell depended on the polyvalent nature of the immune complex. We generated transfectants displaying HEL antigen as a presumptively monovalent integral membrane antigen and used one of these transfectants (J[mHEL]6; Fig. 2a) in excess as a target for HEL-specific B cells. Confocal microscopy showed that, after 10 min, most B cells were in conjunction with a target; BCR was concentrated in the region of synapsis but with clear exclusion of CD45 (Fig. 2b, c, f; Supplementary Information movie 1).

Reorganization of components of the B-cell membrane was also evident from a depletion of CD22 from the centre of most synapses (although often concentrated at the edges), where there was a concentration of ganglioside GM1 (which is associated classically with many Src-family tyrosine kinases). There was also polarization of cytoplasmic components, as judged by a depletion of the signal-inhibitory phosphatase SHP1 in the region of the synapse (Fig. 2d–f), but a concentration of phosphotyrosine-containing proteins as well

as actin and phospholipase C (PLC)- γ 2 (Fig. 2d). This gross reorganization of a B cell after interacting with a target cell may reflect changes that also occur (but possibly on a smaller scale) after BCR crosslinking by soluble antigen, as such crosslinking^{8,9} leads to BCR translocation into a GM1-enriched fraction that is depleted of CD22 and CD45.

We wanted to determine whether reorganization of the B cell could be detected using targets displaying a lower density of a lower-affinity antigen. Transfectants expressing a membrane-integral form of a mutated HEL (mHEL*) that exhibits a 100-fold reduced affinity for the HyHEL10 BCR¹⁰ were screened for low-expressing clones: J[mHEL*]8 displays only around 9×10^3 mHEL* molecules per cell (compared with J[mHEL]6, which expresses about 4×10^5 mHEL per cell).

After incubation of J[mHEL*]8 with HEL-specific B cells, staining for mHEL* (which was previously only dim) was now brightly focused in the synapse (Fig. 3a). This quantity of mHEL* antigen is only sufficient to bring a small proportion of the total BCR into the synapse; however, there is nevertheless a local reorganization of the B cell as the BCR gathering is accompanied by local exclusion of CD45. Notably, the low abundance of mHEL* on the target cell surface is sufficient to trigger B-cell activation, as judged by the upregulation of CD86. Indeed, surface-tethered mHEL* is several orders of magnitude more efficacious than soluble HEL* in this respect (Fig. 3b, c). This is consistent with the apparent increased potency of membrane (as opposed to soluble) antigens in determining B-cell fate *in vivo*^{11,12}, which probably reflects the effective affinity increase resulting from restricting antigen mobility to two dimensions, coupled with the ability to concentrate antigen into the synapse¹³.

To follow synapse formation in real time and exclude the possibility of a fixation artefact, we established targets that displayed mHEL fused to green fluorescent protein (GFP). Incubation with HEL-specific B cells led to rapid aggregation of mHEL–GFP in the synapse (Fig. 3d), although the frequency of synapses decreased at later time points (data not shown). The B cells bind well to their targets even at 13 °C, but formation of a synapse at which BCR and antigen have visibly accumulated requires a higher temperature—consistent with a need for membrane reorganization (Fig. 3e). Similar kinetics of synapse formation were also observed in a different B-cell/target interaction system—that of B cells carrying the 3-83 BCR specific for major histocompatibility complex (MHC) class I molecules with targets that express H2-K^b (Fig. 3f; BCR/antigen affinity estimated¹⁴ to be about 10^5 M⁻¹).

We wondered whether the decrease in observable synapses over time reflected internalization or degradation of antigen–BCR. In T cells, rapid internalization of the antigen receptor is observed after their interaction with antigen-presenting cells¹⁵. After culturing HEL-specific B cells with mHEL targets, flow cytometry showed that there was a rapid downregulation of BCR on the surface of the sorted B cells—more rapid than with soluble antigen (Fig. 4a, b).

This downregulation was peculiar to the BCR and did not apply to the CD19, CD22 or CD45 co-receptors; in addition, experiments performed using B-cell transfectants rather than splenic B cells showed that the downregulation occurring with the antigen-specific BCR did not extend to an irrelevant BCR expressed on the same cell (Fig. 4c). This loss of detectable surface BCR might reflect its internalization by the B cell, its acquisition by the target cell or some form of masking. Immunocytochemistry of cells sorted after co-culture showed that IgM was readily detectable in the sorted B cells (but not the sorted target cells), if they were permeabilized before staining (Fig. 4d). This IgM is unlikely to be *de novo* synthesized IgM that has not yet been transported to the cell surface, as protease stripping of surface IgM on splenic HEL-specific B cells showed that the intracellular pool of IgM was very small. Moreover, co-staining for IgM and HEL showed that, after B-cell–mHEL target interaction, the permeabilized, sorted B cells stained brightly for the

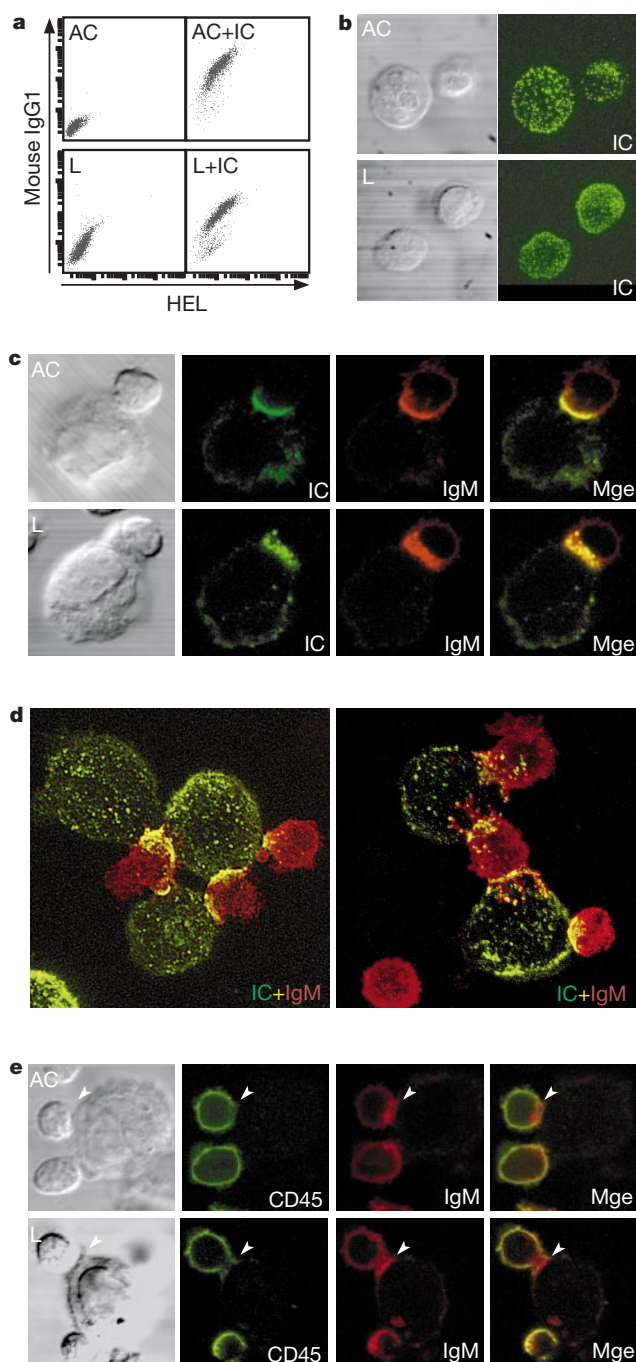


Figure 1 Synapse formation between B cell and antigen-displaying target cell. **a**, Loading of HEL–IgG immune complexes (ICs) onto the 279.AC1 myeloid cell line²⁷ or L-cell transfectants expressing human Fc γ RI. ICs were detected by staining for HEL and IgG1. L, L cells; AC, 279.AC1 cells. **b**, ICs are patchily distributed on the surface of 279.AC1 (top) or L[Fc γ RI] cells (bottom) that have been loaded at 4 °C, fixed and stained for mouse IgG1. Projection views are shown on right. **c**, Synapse formation after incubation (10 min, 37 °C) of IC-loaded targets with excess HEL-specific splenic B cells. For both myeloid and L cells, a single optical section of a single B-cell/target cell interaction is shown in either single or merged view. Whereas most B cells were in contact with a target when IC-loaded L[Fc γ RI] cells were used, only 10% of the B cells formed aggregates with IC-loaded 279.AC1 cells, correlating with internalization of the ICs at 37 °C by the myeloid cells—reflecting their expression of Fc γ RII and Fc γ RIII (not shown). **d**, Projection views of HEL-specific splenic B cells interacting with IC-loaded L[Fc γ RII] cells. Cells were double stained with anti-mouse IgG1 (for ICs; green) and anti-mouse IgM (for BCR; red). **e**, Comparison of BCR (red) and CD45R(B220) (green) distribution on HEL-specific B cells interacting with IC-loaded 279.AC1 or L[Fc γ RII] targets; a single optical section is shown in each case.

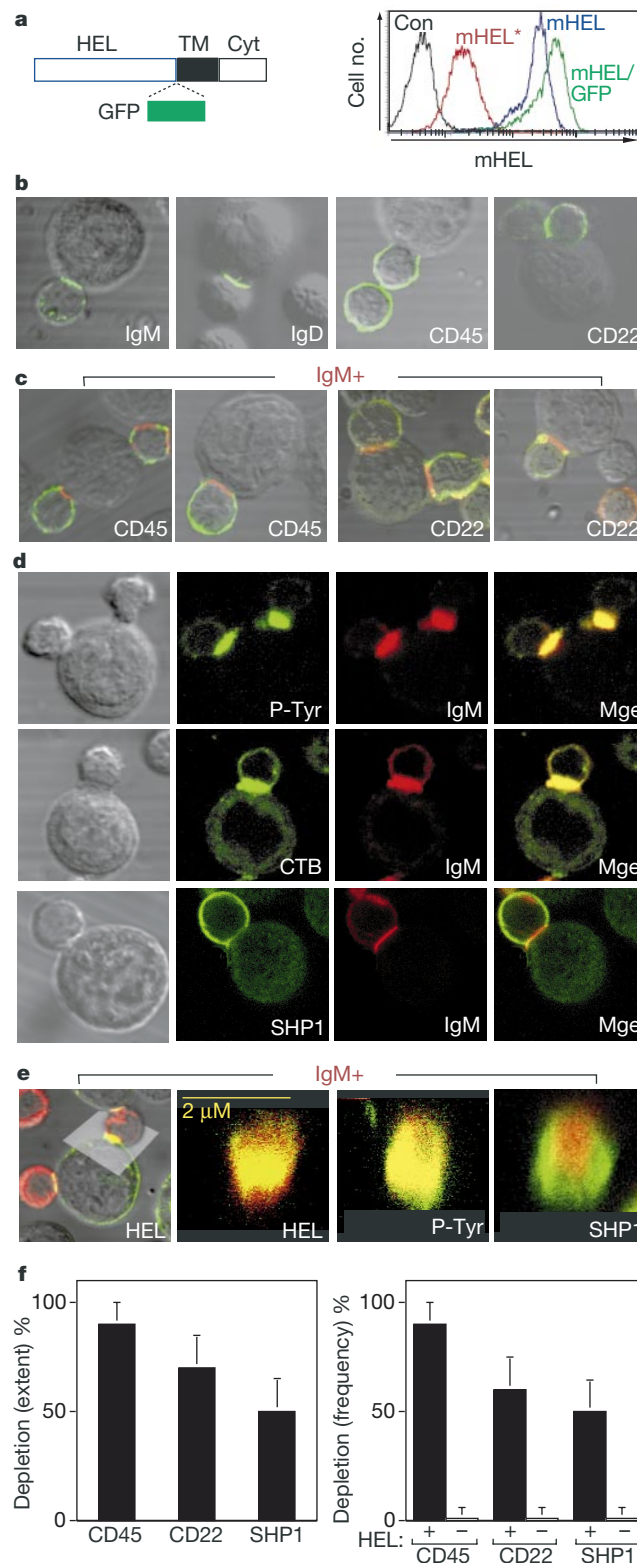


Figure 2 Polarization of the B cell after encountering membrane-immobilized antigen. **a**, Abundance of surface HEL in the cloned J558 transfectants (J[mHEL]6, J[mHEL*]8 and J[mHEL-GFP]) monitored by staining with anti-HEL IgG1- κ (HyHEL5) and PE-conjugated anti- κ . **b**, Distribution of BCR (IgM or IgD), CD22 and CD45 on HEL-specific splenic B cells interacting with J[mHEL]6 targets viewed as single optical sections. Cells were fixed after a 10-min incubation at 37 °C. **c**, Segregation of IgM from CD45 and CD22. Samples were prepared as in **a** but co-stained for CD45 or CD22 (green) and IgM (red). **d**, Polarization of HEL-specific B cells interacting with J[mHEL]6 targets analysed by double staining, in each case showing a single confocal section in single or merged view. Ganglioside GM1 was detected using cholera toxin B subunit (CTB). See Supplementary Information for

additional staining for actin and PLC- γ 2. **e**, Views along the plane of synopsis co-staining for IgM (red) and HEL, phosphotyrosine or SHP1 (green), showing coincidence of HEL-IgM and phosphotyrosine-IgM at the synapse but exclusion of SHP1. **f**, Quantification of the depletion of CD45, CD22 and SHP1 at synapses. Left, the mean extent of depletion of the appropriate marker at the synapse (from analysis of an average of 50 synapses) is expressed relative to the abundance of that marker adjacent to the synapse but in the same confocal plane. Right, the percentage of antigen-specific synapses ('HEL+') in which such depletion was noted (CD45, below 80%; CD22 and SHP1, below 60%) is compared with depletion in contacts between B cells and antigen-lacking targets ('HEL-').

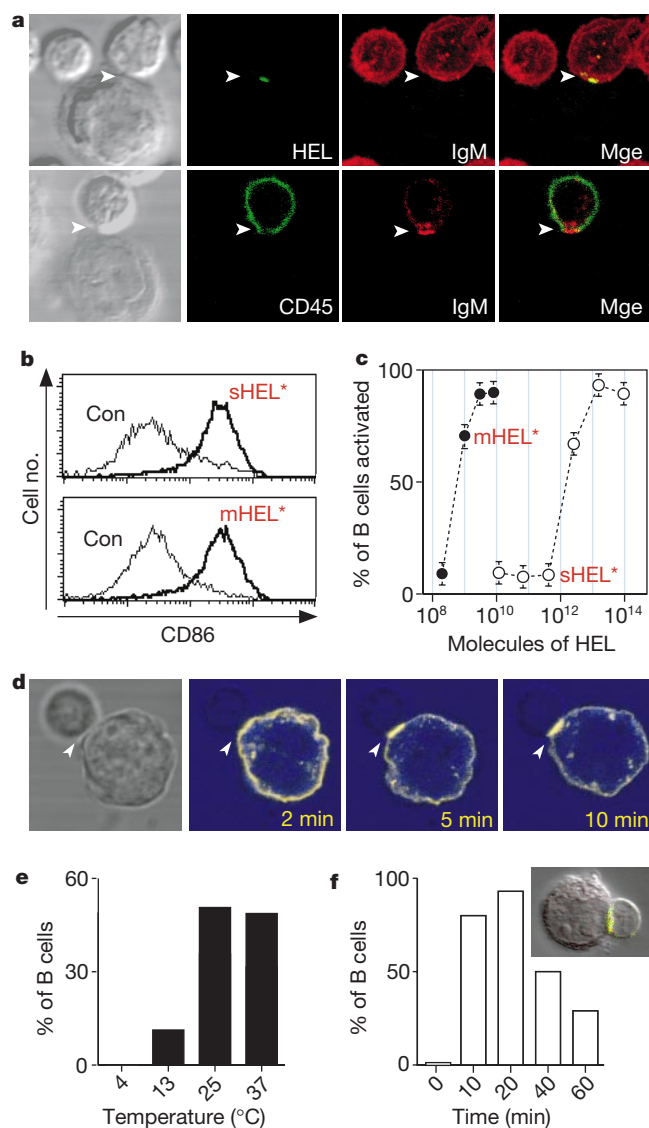


Figure 3 Parameters affecting synapse formation. **a**, Synapse formation between HEL-specific splenic B cells and J[mHEL*]8 targets (see Fig. 2a), which express antigen of reduced affinity at low density. Top, single or merged view of cells stained for HEL and IgM. Bottom, staining for CD45 and IgM, showing a single optical section. **b**, Culture of HEL-specific splenic B cells (20 h) with either J[mHEL*]8 (twofold excess) or soluble HEL ($0.1 \mu\text{g ml}^{-1}$) causes similar B-cell activation as judged by upregulation of CD86. **c**, Comparison of B-cell activation by membrane-tethered and soluble HEL. HEL-specific splenic B cells were cultured with varying concentrations of soluble HEL* or varying numbers of J558[mHEL*] targets, and activation was monitored as the percentage of B cells with upregulated expression of CD86. **d**, Real-time microscopy of an HEL-specific splenic B cell interacting at 20°C with an mHEL–GFP-expressing target. **e**, Temperature dependence of synapse formation, as judged by the percentage of B cells bound to a target in which 50% of the IgM in one optical section was concentrated in the region of synapsis. **f**, Synapse frequency as a function of time. The experiment shown (as in **e** but at 37°C) was performed using H2K^b-specific splenic B cells from 3–83 mice and H2K^b hybridoma targets. Shown is a single-plane confocal image of such a synapse (co-stained for IgM, red; H2K^b, green).

antigen, indicating that downregulation of the BCR might reflect acquisition of the antigen from the target cell and internalization of the mHEL–BCR complex by the B cell (Fig. 4e).

If the B cell acquires the antigen from the target cell, this should be visible in living cells using mHEL–GFP targets, without the risk of flow cytometry or fixation-associated artefacts. This is indeed the case. Following the interaction at 20°C in real time shows that

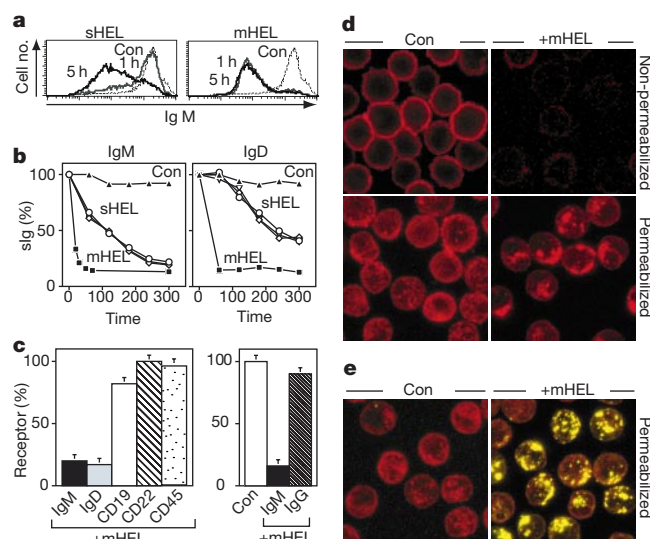


Figure 4 Downregulation of surface BCR after antigen encounter. **a**, Downregulation of IgM on HEL-specific splenic B cells after incubation (1 or 5 h at 37°C) with saturating ($10 \mu\text{g ml}^{-1}$) amounts of soluble HEL or J[mHEL]6 target cells. Thin line indicates unstimulated cells. After incubation, cells were analysed by flow cytometry, gating on scatter and CD45R(B220) expression. **b**, Surface IgM and IgD expression on HEL-specific splenic B cells as a function of incubation time at 37°C with J[mHEL]6 targets (filled boxes), control targets (untransfected J558; filled triangles) or soluble HEL (at 10, 1 and $0.1 \mu\text{g ml}^{-1}$; open symbols). Surface immunoglobulin (slg) expression at each time point is presented as the percentage mean surface immunoglobulin fluorescence relative to that of untreated cells. **c**, Antigen-specific BCR and not CD19, CD22, CD45 or an irrelevant BCR is downregulated. Left, expression of the indicated markers was monitored on HEL-specific splenic B cells after a 1-h incubation at 37°C with J[mHEL]6 targets. Right, surface IgM and IgG2a expression was monitored in a parallel experiment but using mouse B-lymphoma transfectants³⁰ that co-express an HEL-specific IgM and a non-HEL-specific IgG2a BCR. Con, IgM expression after with incubation with untransfected J558 targets. **d**, Downregulated surface IgM can be revealed by permeabilization treatment of the B cell. HEL-specific splenic B cells were incubated at 37°C for 1 h with J[mHEL]6 targets (or untransfected J558 controls, left), sorted on the basis of scatter and CD45R(B220) expression, fixed, and subjected or not to permeabilization treatment before staining with anti-IgM (red) and analysis by immunofluorescence microscopy. **e**, HEL colocalizes with downregulated IgM in the B cell. Samples as in **d** but stained for HEL (green) and IgM (red).

synapse formation is followed by the appearance of GFP fluorescence at points in the B cell that are distant from the synapse (Fig. 5a); this effect is more marked if the cells are incubated at 37°C (Fig. 5b). Furthermore, if these cells are fixed and permeabilized, antigen transfer is evident from the fact that serial sections reveal co-localization of mHEL–GFP and IgM at sites in the B cell that are both removed from, and in a different plane to that of the synapse (Fig. 5c; Supplementary Information movie 2). The mHEL–GFP that has transferred to the B cell seems to undergo degradation, as the GFP fluorescence of the B cell diminishes with time (Fig. 5d).

The ability of the B cell to acquire an integral membrane antigen from the target cell and then internalize and degrade it should presumably allow the B cell to present antigen-derived peptides to T cells. HEL-specific B-cell transfectants co-cultured with MHC class-II-negative mHEL targets are very effective in presenting HEL-derived peptides to T cells (Fig. 5e). This presentation does not simply reflect BCR-mediated uptake of soluble mHEL fragments that have been released spontaneously by the target cells, as presentation is abolished if the B cells and target cells are separated by a mesh.

This presentation of membrane-tethered antigens is impressively efficient: co-incubation of 10^5 HEL-specific B cells with only $3 \times$

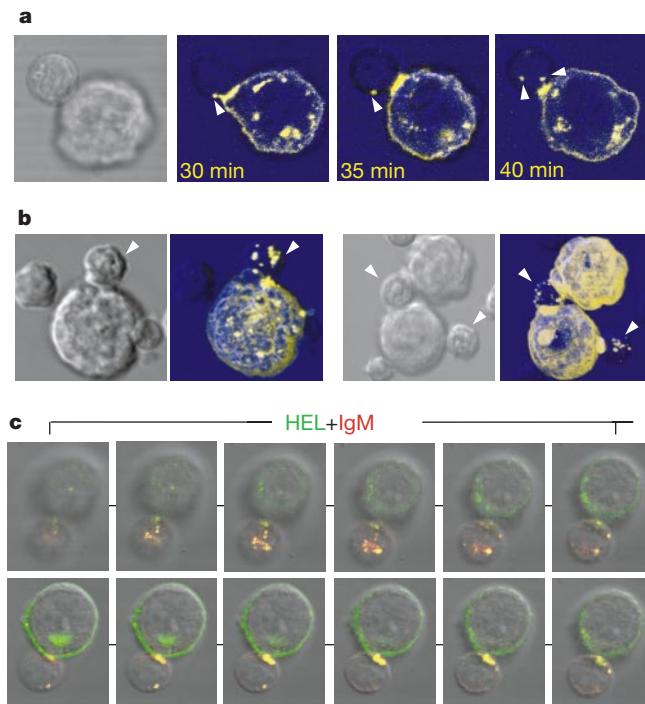
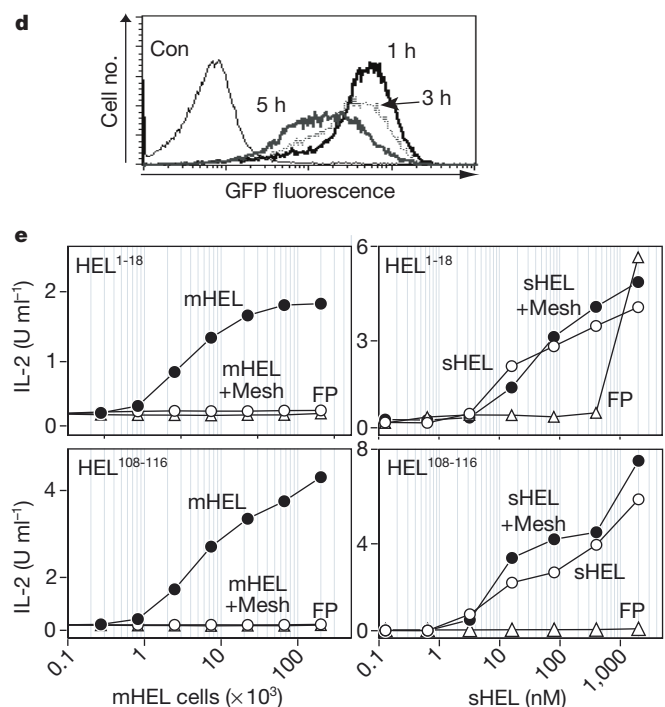


Figure 5 B cells acquire and present integral membrane antigens from the target cell. **a**, Real-time microscopy (single optical section) of an HEL-specific splenic B cell interacting with an mHEL-GFP target at 20 °C. **b**, Two examples of antigen acquisition from an mHEL-GFP target showing live cells (projection views) after a 40-min incubation at 37 °C. **c**, Serial optical sections of a fixed and permeabilized HEL-specific splenic B cell (smaller cell) interacting with a J[mHEL]6 target, showing HEL (green)/BCR (red) colocalization at sites in the B cell distant from the synapse. **d**, GFP fluorescence diminishes with time in HEL-specific splenic B cells incubated with J[mHEL/GFP] targets. GFP fluorescence in scatter and B220-gated B cells was monitored after 1, 3 and 5 h of



incubation with targets. **e**, Presentation of HEL-derived epitopes to T cells by HEL-specific B-cell transfectants incubated with either J[mHEL]6 targets (left) or soluble HEL (right). B cells (10^5) and target cells (or soluble HEL, as indicated) were cultured (0.3 ml) with T-cell hybridomas specific for HEL¹⁻¹⁸ or HEL¹⁰⁸⁻¹¹⁶ in the context of MHC class II molecules with, in some cases (open symbols), the target cells separated from the B cell/T cell mix by a mesh (0.2-μm Anopore membrane; Nunc). T-cell activation was monitored by measuring IL-2 production after 24 h. FP, fluid phase presentation by B cells lacking an HEL-specific BCR.

10^3 mHEL-expressing target cells (which correspond to a total of some 10^8 – 10^9 surface-expressed mHEL molecules) is sufficient to yield detectable interleukin (IL)-2 production in the antigen-presentation assay. In contrast, presentation of soluble HEL under similar assay conditions requires more than 10^{11} molecules of antigen. The increased efficacy of presentation of tethered (as opposed to soluble) antigen is even more striking when using mutated HELs exhibiting lower affinity for the BCR (data not shown).

The precise nature of the synapse is at present unclear, although co-staining with clathrin and early endosomal antigen suggests that much of the BCR-antigen complex may be internalized at the synapse (data not shown). Nevertheless, the synapse formed between the B cell and the antigen-displaying target cell shares some obvious similarities to the synapse formed between T cells and their targets, such as a concentration of antigen receptor but exclusion of CD45 from the region of synapsis^{16–19}. But whereas T cells recognize antigen displayed in the context of MHC on many autologous cell types, it would seem beneficial if B-cell activation by antigen displayed on autologous cells were biased towards antigens tethered by professional receptors (such as FcR or CD21) and against abundant membrane antigens of self. One might imagine that there are surface glycoproteins on autologous cells that mediate selective recruitment of appropriate BCR co-receptors into (or out of) the synapse and thereby bias B-cell activation to FcR/CD21-tethered antigens.

The BCR-mediated concentration of antigen into the synapse is followed by antigen acquisition. Acquisition of an integral membrane protein from a target cell is not without precedent^{20–22}, although the mechanism remains unknown. Proteolytic release of

the antigen from the target remains a possibility, but our failure to detect such proteolysis²³ leads us to favour a BCR-mediated pinching of membrane vesicles from the target cell. Our results show that the acquisition allows extremely efficient antigen presentation, which probably not only reflects the actual acquisition of the antigen: the gathering of BCR and the resultant polarization of the B cell may also enhance the antigen processing machinery of the B cell^{24–26}. Together with the fact that membrane tethering of antigen also leads to great enhancement of B-cell signalling, as judged by the upregulation of CD86 (Fig. 3c), it seems that the focusing of antigen into the synapse may be important in B-cell responses to antigen—especially at low antigen abundance. With respect to membrane-integral antigens on foreign target cells, the gathering of antigen into the synapse and its consequent acquisition will presumably enhance the extent to which cognate B cells will present peptides to T cells that are derived from proteins linked to the antigen recognized by the BCR (rather than from irrelevant proteins in the foreign target). This might favour efficient recruitment of T-cell help and bias against promiscuous T-cell activation and the development of autoimmunity. □

Methods

Loading and detection of immune complexes

To load immune complexes onto 279.AC1 myeloid cells²⁷ or L-cell transfectants expressing human FcγRI (a gift from M. Clark), we incubated cells as described²³ with a mixture of biotinylated-HEL, HEL-specific IgG1 HyHEL5 and F10 (which recognize HEL epitopes distinct from that recognized by the HyHEL10 BCR expressed by B cells from MD4 transgenic mice) and rabbit anti-mouse IgG. Loaded immune complexes were detected using fluorescein isothiocyanate (FITC)-streptavidin and PE-conjugated goat anti-mouse IgG1. We detected SHP1 and phosphorylated tyrosine (P-Tyr) using rabbit polyclonal antisera (SantaCruz) and biotinylated monoclonal antibody 4G10 (Upstate Biotechnol-

ogy), respectively. Staining specificity was controlled by single staining, as well as by using secondary antibodies in the absence of the primary stain.

Generation of target cells

Target cells displaying a membrane-integral version of either wild-type HEL or a mutant¹⁰ exhibiting reduced affinity for HyHEL10 ([R²¹, D¹⁰¹, G¹⁰², N¹⁰³] designated HEL*) were generated by transfecting mouse J558L plasmacytoma cells with constructs analogous to those used¹⁰ for expression of soluble HEL/HEL*, except that 14 Ser/Gly codons, the H2K^b transmembrane region, and a 23-codon cytoplasmic domain were inserted immediately upstream of the termination codon by polymerase chain reaction. For mHEL-GFP, we included the EGFP coding domain in the Ser/Gly linker. Abundance of surface HEL was monitored by flow cytometry and radiolabelled-antibody binding using HyHEL5 and D1.3 HEL-specific monoclonal antibodies, for which the mutant HELs used in this work show unaltered affinities¹⁰.

Interaction assays

For B-cell/target interaction assays, splenic B cells from 3-83 or MD4 transgenic mice^{28,29} carrying (IgM + IgD) BCRs specific for HEL or H2K^b/H2K^b were freshly purified on Lympholyte and incubated with a twofold excess of target cells in RPMI, 50 mM HEPES pH 7.4, for the appropriate time at 37 °C before being applied to polylysine-coated slides. Cells were fixed in 4% paraformaldehyde/PBS or methanol and permeabilized with PBS/0.1% Triton X-100 before immunofluorescence. We acquired confocal images using a Nikon E800 microscope attached to BioRad Radiance Plus scanning system equipped with 488-nm and 543-nm lasers, as well as differential interference contrast for transmitted light. GFP fluorescence in living cells in real time was visualized using a Radiance 2000 and Nikon E300 inverted microscope. Images were processed using BioRad Lasersharp 1024 or 2000 software to provide single plane images, confocal projections or slicing.

Antigen presentation

Presentation of HEL epitopes to T-cell hybridomas 2G7 (specific for I-E^k[HEL¹⁻¹⁸]) and 1E5 (specific for I-E^d[HEL¹⁰⁸⁻¹¹⁶]) by transfectants of the LK35.2 B-cell hybridoma expressing an HEL-specific IgM BCR was monitored as described¹⁰.

Received 12 December 2000; accepted 30 March 2001.

- Lanzavecchia, A. Antigen-specific interaction between T and B cells. *Nature* **314**, 537–539 (1985).
- Klaus, G. G., Humphrey, J. H., Kunkl, A. & Dongworth, D. W. The follicular dendritic cell: its role in antigen presentation in the generation of immunological memory. *Immunol. Rev.* **53**, 3–28 (1980).
- Tew, J. G., Kosco, M. H., Burton, G. F. & Szakal, A. K. Follicular dendritic cells as accessory cells. *Immunol. Rev.* **117**, 185–211 (1990).
- Kosco-Vilbois, M. H., Gray, D., Scheidegger, D. & Julius, M. Follicular dendritic cells help resting B cells to become effective antigen-presenting cells: induction of B7/BB1 and upregulation of major histocompatibility complex class II molecules. *J. Exp. Med.* **178**, 2055–2066 (1993).
- Schamel, W. W. & Reth, M. Monomeric and oligomeric complexes of the B cell antigen receptor. *Immunity* **13**, 5–14 (2000).
- Taylor, R. B., Duffus, W. P. H., Raff, M. C. & de Petris, S. Redistribution and pinocytosis of lymphocyte surface immunoglobulin molecules induced by anti-immunoglobulin antibody. *Nature* **233**, 225–227 (1971).
- Schreiner, G. F. & Unanue, E. R. Capping and the lymphocyte: models for membrane reorganization. *J. Immunol.* **119**, 1549–1551 (1977).
- Cheng, P. C., Dykstra, M. L., Mitchell, R. N. & Pierce, S. K. A role for lipid rafts in B cell antigen receptor signaling and antigen targeting. *J. Exp. Med.* **190**, 1549–1560 (1999).
- Weintraub, B. C. *et al.* Entry of B cell receptor into signaling domains is inhibited in tolerant B cells. *J. Exp. Med.* **191**, 1443–1448 (2000).
- Batista, F. D. & Neuberger, M. S. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* **8**, 751–759 (1998).
- Nemazee, D. & Burki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* **337**, 562–566 (1989).
- Hartley, S. B. *et al.* Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature* **353**, 765–769 (1991).
- Dustin, M. L. *et al.* Low affinity interaction of human and rat T cell adhesion molecule CD2 with its ligands aligns adhering membranes to achieve high physiological affinity. *J. Biol. Chem.* **272**, 30889–30898 (1997).
- Lang, J. *et al.* B cells are exquisitely sensitive to central tolerance and receptor editing by ultralow affinity, membrane-bound antigen. *J. Exp. Med.* **184**, 1685–1697 (1996).
- Valitutti, S., Muller, S., Cella, M., Padovan, E. & Lanzavecchia, A. Serial triggering of many T-cell receptors by a few peptide–MHC complexes. *Nature* **375**, 148–151 (1995).
- Monks, C. R., Freiberg, B. A., Kupfer, H., Sciaky, N. & Kupfer, A. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* **395**, 82–86 (1998).
- Wulfig, C. & Davis, M. M. A receptor/cytoskeletal movement triggered by costimulation during T cell activation. *Science* **282**, 2266–2269 (1998).
- Grakoui, A. *et al.* The immunological synapse: a molecular machine controlling T cell activation. *Science* **285**, 221–227 (1999).
- Leupin, O., Zaru, R., Laroche, T., Muller, S. & Valitutti, S. Exclusion of CD45 from the T-cell receptor signaling area in antigen-stimulated T lymphocytes. *Curr. Biol.* **10**, 277–280 (2000).
- Cagan, R. L., Kramer, H., Hart, A. C. & Zipursky, S. L. The bride of sevenless and sevenless interaction: internalization of a transmembrane ligand. *Cell* **69**, 393–399 (1992).
- Huang, J. F. *et al.* TCR-mediated internalization of peptide–MHC complexes acquired by T cells. *Science* **286**, 952–954 (1999).
- Hwang, I. *et al.* T cells can use either T cell receptor or CD28 receptors to absorb and internalize cell surface molecules derived from antigen-presenting cells. *J. Exp. Med.* **191**, 1137–1148 (2000).
- Batista, F. D. & Neuberger, M. S. B cells extract and present immobilized antigen: implications for affinity discrimination. *EMBO J.* **19**, 513–520 (2000).
- Casten, L. A., Lakey, E. K., Jelachich, M. L., Margoliash, E. & Pierce, S. K. Anti-immunoglobulin

- augments the B-cell antigen-presentation function independently of internalization of receptor-antigen complex. *Proc. Natl Acad. Sci. USA* **82**, 5890–5894 (1985).
- Siemasko, K., Eisfelder, B. J., Williamson, E., Kabak, S. & Clark, M. R. Signals from the B lymphocyte antigen receptor regulate MHC class II containing late endosomes. *J. Immunol.* **160**, 5203–5208 (1998).
- Serre, K. *et al.* Efficient presentation of multivalent antigens targeted to various cell surface molecules of dendritic cells and surface Ig of antigen-specific B cells. *J. Immunol.* **161**, 6059–6067 (1998).
- Green, S. M., Lowe, A. D., Parrington, J. & Karn, J. Transformation of growth factor-dependent myeloid stem cells with retroviral vectors carrying c-myc. *Oncogene* **737**–751 (1989).
- Russell, D. M. *et al.* Peripheral deletion of self-reactive B cells. *Nature* **354**, 308–311 (1991).
- Goodnow, C. C. *et al.* Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* **334**, 676–682 (1988).
- Aluvihare, V. R., Khamlichi, A. A., Williams, G. T., Adorini, L. & Neuberger, M. S. Acceleration of intracellular targeting of antigen by the B-cell antigen receptor: importance depends on the nature of the antigen-antibody interaction. *EMBO J.* **16**, 3553–3562 (1997).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank B. Amos and S. Reichelt for help and advice with confocal microscopy, and S. Munro for helpful discussions. We are indebted those who provided antibodies, transgenic mice and cell lines. F.D.B. and D.I. were supported by the Arthritis Research Campaign and Studienstiftung des deutschen Volkes, respectively.

Correspondence and requests for materials should be addressed to F.D.B. (e-mail: fdb@mrc-lmb.cam.ac.uk) or M.S.N. (e-mail: msn@mrc-lmb.cam.ac.uk)

Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells

Sayda M. Elbashir*, Jens Harborth†, Winfried Lendeckel*, Abdullah Yalcin*, Klaus Weber† & Thomas Tuschl*

* Department of Cellular Biochemistry; and † Department of Biochemistry and Cell Biology, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

RNA interference (RNAi) is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene^{1–4}. The mediators of sequence-specific messenger RNA degradation are 21- and 22-nucleotide small interfering RNAs (siRNAs) generated by ribonuclease III cleavage from longer dsRNAs^{5–9}. Here we show that 21-nucleotide siRNA duplexes specifically suppress expression of endogenous and heterologous genes in different mammalian cell lines, including human embryonic kidney (293) and HeLa cells. Therefore, 21-nucleotide siRNA duplexes provide a new tool for studying gene function in mammalian cells and may eventually be used as gene-specific therapeutics.

Uptake of dsRNA by insect cell lines has previously been shown to 'knock-down' the expression of specific proteins, owing to sequence-specific, dsRNA-mediated mRNA degradation^{6,10–12}. However, it has not been possible to detect potent and specific RNA interference in commonly used mammalian cell culture systems, including 293 (human embryonic kidney), NIH/3T3 (mouse fibroblast), BHK-21 (Syrian baby hamster kidney), and CHO-K1 (Chinese hamster ovary) cells, applying dsRNA that varies in size between 38 and 1,662 base pairs (bp)^{10,12}. This apparent lack of RNAi in mammalian cell culture was unexpected, because RNAi exists in mouse oocytes and early embryos^{13,14}, and because RNAi-related, transgene-mediated co-suppression was also observed in cultured Rat-1 fibroblasts¹⁵. But it is known that dsRNA in the cytoplasm of mammalian cells can trigger profound physiological

reactions that lead to the induction of interferon synthesis¹⁶. In the interferon response, dsRNA > 30 bp binds and activates the protein kinase PKR¹⁷ and 2',5'-oligoadenylate synthetase (2',5'-AS)¹⁸. Activated PKR stalls translation by phosphorylation of the translation initiation factors eIF2 α , and activated 2',5'-AS causes mRNA degradation by 2',5'-oligoadenylate-activated ribonuclease L. These responses are intrinsically sequence-nonspecific to the inducing dsRNA.

Base-paired 21- and 22-nucleotide (nt) siRNAs with overhanging 3' ends mediate efficient sequence-specific mRNA degradation in lysates prepared from *Drosophila* embryos⁹. To test whether siRNAs are also capable of mediating RNAi in cell culture, we synthesized 21-nt siRNA duplexes with symmetric 2-nt 3' overhangs directed against reporter genes coding for sea pansy (*Renilla reniformis*, RL) and two sequence variants of firefly (*Photinus pyralis*, GL2 and GL3) luciferases (Fig. 1a, b). The siRNA duplexes were co-transfected with the reporter plasmid combinations pGL2/pRL or pGL3/pRL, into *Drosophila* S2 cells or mammalian cells using cationic liposomes. Luciferase activities were determined 20 h after transfection. In *Drosophila* S2 cells (Fig. 2a and b), the specific inhibition of luciferases was complete and similar to results previously obtained for longer dsRNAs^{6,10,12,19}. In mammalian cells, where the reporter genes were 50- to 100-fold more strongly expressed, the specific suppression was less complete (Fig. 2c-j). In NIH/3T3, monkey COS-7 and HeLa S3 cells (Fig. 2c-h), GL2 expression was reduced 3-

to 12-fold, GL3 expression 9- to 25-fold, and RL expression 2- to 3-fold, in response to the cognate siRNAs. For 293 cells, targeting of RL luciferase by RL siRNAs was ineffective, although GL2 and GL3 targets responded specifically (Fig. 2i and j). The lack of reduction of RL expression in 293 cells may be because of its expression, 5- to 20-fold higher than any other mammalian cell line tested and/or to limited accessibility of the target sequence due to RNA secondary structure or associated proteins. Nevertheless, specific targeting of GL2 and GL3 luciferase by the cognate siRNA duplexes indicated that RNAi is also functioning in 293 cells.

The 2-nucleotide 3' overhang in all siRNA duplexes was composed of (2'-deoxy) thymidine, except for uGL2, which contained

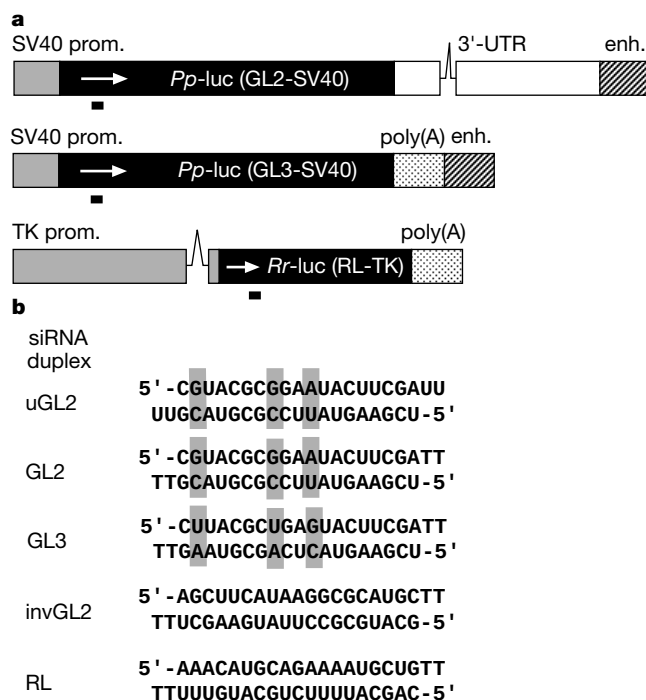


Figure 1 Reporter constructs and siRNA duplexes. **a**, The firefly (*Pp-luc*) and sea pansy (*Rr-luc*) luciferase reporter-gene regions from plasmids pGL2-Control, pGL3-Control, and pRL-TK (Promega) are illustrated; simian virus 40 (SV40) promoter (prom.); SV40 enhancer element (enh.); SV40 late polyadenylation signal (poly(A)); herpes simplex virus (HSV) thymidine kinase promoter, and two introns (lines) are indicated. The sequence of GL3 luciferase is 95% identical to GL2, but RL is completely unrelated to both. Luciferase expression from pGL2 is approximately 10-fold lower than from pGL3 in transfected mammalian cells. The region targeted by the siRNA duplexes is indicated as black bar below the coding region of the luciferase genes. **b**, The sense (top) and antisense (bottom) sequences of the siRNA duplexes targeting GL2, GL3, and RL luciferase are shown. The GL2 and GL3 siRNA duplexes differ by only three single-nucleotide substitutions (boxed in grey). As nonspecific control, a duplex with the inverted GL2 sequence, invGL2, was synthesized. The 2-nucleotide 3' overhang of 2'-deoxythymidine is indicated as TT; uGL2 is similar to GL2 siRNA but contains ribo-uridine 3' overhangs.

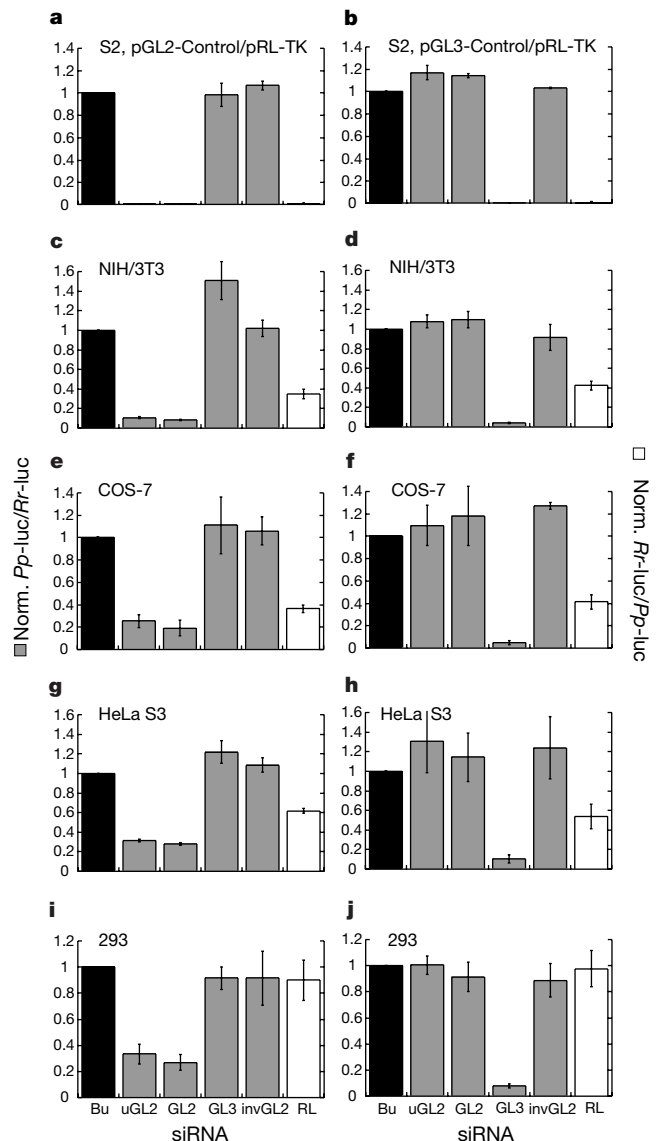


Figure 2 RNA interference by siRNA duplexes. Ratios of target to control luciferase were normalized to a buffer control (Bu, black bars); grey bars indicate ratios of *Photinus pyralis* (*Pp-luc*) GL2 or GL3 luciferase to *Renilla reniformis* (*Rr-luc*) RL luciferase (left axis), white bars indicate RL to GL2 or GL3 ratios (right axis). **a, c, e, g** and **i**, Experiments performed with the combination of pGL2-Control and pRL-TK reporter plasmids; **b, d, f, h** and **j**, experiments performed with the combination of pGL3-Control and pRL-TK reporter plasmids. The cell line used for the interference experiment is indicated at the top of each plot. The ratios of *Pp-luc*/*Rr-luc* for the buffer control (Bu) varied between 0.5 and 10 for pGL2/pRL, and between 0.03 and 1 for pGL3/pRL, respectively, before normalization and between the various cell lines tested. The plotted data were averaged from three independent experiments $\pm$ s.d.

uridine residues. The thymidine overhang was chosen because it reduces costs of RNA synthesis and may enhance nuclease resistance of siRNAs in the cell culture medium and within transfected cells. As in the *Drosophila in vitro* system (data not shown), substitution of uridine by thymidine in the 3' overhang was well tolerated in cultured mammalian cells (Fig. 2a, c, e, g and i), and the sequence of the overhang appears not to contribute to target recognition⁹.

In co-transfection experiments, 25 nM siRNA duplexes were used (Figs 2 and 3; concentration is in respect to the final volume of tissue culture medium). Increasing the siRNA concentration to 100 nM did not enhance the specific silencing effects, but started to affect transfection efficiencies, perhaps due to competition for liposome encapsulation between plasmid DNA and siRNA (data not shown). Decreasing the siRNA concentration to 1.5 nM did not reduce the specific silencing effect (data not shown), even though the siRNAs were now only 2- to 20-fold more concentrated than the DNA plasmids; the silencing effect only vanishes completely if the siRNA concentration was dropped below 0.05 nM. This indicates that siRNAs are extraordinarily powerful reagents for mediating gene silencing, and that siRNAs are effective at concentrations that are several orders of magnitude below the concentrations applied in conventional antisense or ribozyme gene-targeting experiments²⁰.

To monitor the effect of longer dsRNAs on mammalian cells, 50- and 500-bp dsRNAs that are cognate to the reporter genes were prepared. As a control for nonspecific inhibition, dsRNAs from humanized GFP (hG)²¹ was used. In these experiments, the reporter plasmids were co-transfected with either 0.21 µg siRNA duplexes or 0.21 µg longer dsRNAs. The siRNA duplexes only reduced the expression of their cognate reporter gene, while the longer dsRNAs strongly and nonspecifically reduced reporter-gene expression. The effects are illustrated for HeLa S3 cells as a representative example (Fig. 3a and b). The absolute luciferase activities were decreased nonspecifically 10- to 20-fold by 50-bp dsRNA, and 20- to 200-fold by 500-bp dsRNA co-transfection, respectively. Similar nonspecific effects were observed for COS-7 and NIH/3T3 cells. For 293 cells, a 10- to 20-fold nonspecific reduction was observed only for 500-bp dsRNAs. Nonspecific reduction in reporter-gene expression by dsRNA > 30 bp was expected as part of the interferon response¹⁶. Interestingly, superimposed on the nonspecific interferon response, we detect additional sequence-specific, dsRNA-mediated silencing. The sequence-specific silencing effect of long dsRNAs, however, became apparent only when the relative reporter-gene activities were normalized to the hG dsRNA controls (Fig. 3c). Sequence-specific silencing by 50- or 500-bp dsRNAs reduced the targeted reporter-gene expression by an additional 2- to 5-fold. Similar effects were also detected in the other three mammalian cell lines tested (data not shown). Specific silencing effects with dsRNAs (356–1,662 bp) were previously reported in CHO-K1 cells, but the amounts of dsRNA required to detect a 2- to 4-fold specific reduction were about 20-fold higher than in our experiments¹². Also, CHO-K1 cells appear to be deficient in the interferon response. In another report, 293, NIH/3T3 and BHK-21 cells were tested for RNAi using luciferase/β-galactosidase (lacZ) reporter combinations and 829-bp specific lacZ or 717-bp nonspecific green fluorescent protein (GFP) dsRNA¹⁰. The lack of detected RNAi in this case may be due to the less sensitive luciferase/lacZ reporter assay and the length differences of target and control dsRNA. Taken together, our results indicate that RNAi is active in mammalian cells, but that the silencing effect is difficult to detect if the interferon system is activated by dsRNA > 30 bp.

To test for silencing of endogenous genes, we chose four genes coding for cytoskeletal proteins: lamin A/C, lamin B1, nuclear mitotic apparatus protein (NuMA) and vimentin²⁷. The selection was based on the availability of antibodies needed to quantitate the silencing effect. Silencing was monitored 40 to 45 h after transfection to allow for turnover of the protein of the targeted genes. As

shown in Fig. 4, the expression of lamin A/C was specifically reduced by the cognate siRNA duplex (Fig. 4a), but not when nonspecific siRNA directed against firefly luciferase (Fig. 4b) or buffer (Fig. 4c) was used. The expression of a non-targeted gene, NuMA, was unaffected in all treated cells (Fig. 4d–f), demonstrating the integrity of the targeted cells. The reduction in lamin A/C proteins was more than 90% complete as quantified by western blotting (Fig. 4j, k). We note that lamin A/C 'knock-out' mice are

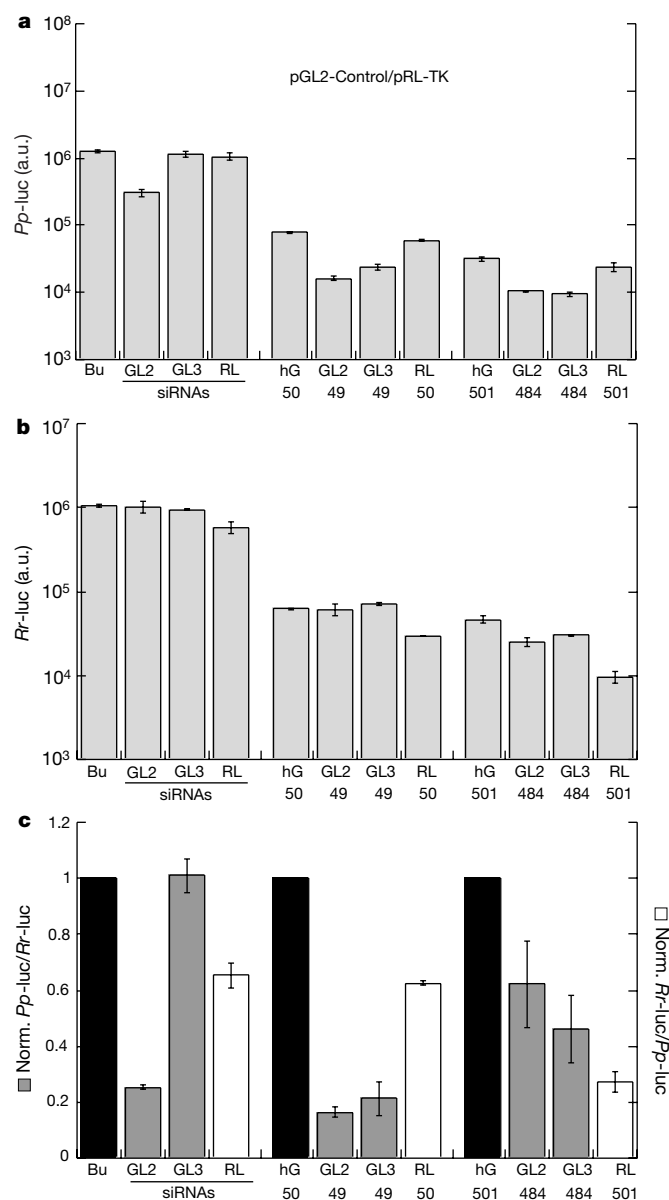


Figure 3 Effects of 21-nucleotide siRNAs, 50-bp, and 500-bp dsRNAs on luciferase expression in HeLa cells. The exact length of the long dsRNAs in base pairs is indicated below the bars. Experiments were performed with pGL2-Control and pRL-TK reporter plasmids. The data were averaged from two independent experiments $\pm$ s.d. **a**, Absolute Pp-luc expression, plotted in arbitrary luminescence units (a.u.). **b**, Rr-luc expression, plotted in arbitrary luminescence units. **c**, Ratios of normalized target to control luciferase. The ratios of luciferase activity for siRNA duplexes were normalized to a buffer control (Bu, black bars); the luminescence ratios for 50- or 500-bp dsRNAs were normalized to the respective ratios observed for 50- and 500-bp dsRNAs from humanized GFP (hG, black bars). We note that the overall differences in sequence between the 49- and 484-bp GL2 and GL3 dsRNAs are not sufficient to confer specificity for targeting GL2 and GL3 targets (43-nucleotide uninterrupted identity in 49-bp segment, 239-nucleotide longest uninterrupted identity in 484-bp segment)³⁰.

viable for a few weeks after birth²³ and that the lamin A/C knock-down in cultured cells was not expected to cause cell death. Lamin A and C are produced by alternative splicing in the 3' region and are present in equal amounts in the lamina of mammalian cells (Fig. 4j, k). Transfection of siRNA duplexes targeting lamin B1 and NuMA reduced the expression of these proteins to low levels (data not shown), but we were not able to observe a reduction in vimentin expression. This could be due to the high abundance of vimentin in the cells (several per cent of total cell mass) or because the siRNA sequence chosen was not optimal for targeting of vimentin.

The mechanism of the 21-nucleotide siRNA-mediated interference process in mammalian cells remains to be uncovered, and silencing might occur post-transcriptionally and/or transcriptionally. In *Drosophila* lysate, siRNA duplexes mediate post-transcriptional gene silencing by reconstitution of siRNA-protein complexes (siRNPs), which guide mRNA recognition and targeted cleavage^{6,7,9}. In plants, dsRNA-mediated post-transcriptional silencing has also been linked to DNA methylation, which may also be directed by 21-

nucleotide siRNAs²⁴. Methylation of promoter regions can lead to transcriptional silencing²⁵, but methylation in coding sequences does not²⁶. DNA methylation and transcriptional silencing in mammals are well documented processes²⁷, yet their mechanisms have not been linked to that of post-transcriptional silencing. Methylation in mammals is predominantly directed towards CpG dinucleotide sequences. There is no CpG sequence in the RL or lamin A/C siRNA, although both siRNAs mediate specific silencing in mammalian cell culture, so it is unlikely that DNA methylation is essential for the silencing process.

Thus we have shown, for the first time, siRNA-mediated gene silencing in mammalian cells. The use of exogenous 21-nucleotide siRNAs holds great promise for analysis of gene function in human cell culture and the development of gene-specific therapeutics. It will also be of interest in understanding the potential role of endogenous siRNAs in the regulation of mammalian gene function. □

Methods

RNA preparation

21-nucleotide RNAs were chemically synthesized using Expedite RNA phosphoramidites and thymidine phosphoramidite (Prologo, Germany). Synthetic oligonucleotides were deprotected and gel-purified⁹. The accession numbers given below are from GenBank. The siRNA sequences targeting GL2 (Acc. No. X65324) and GL3 luciferase (Acc. No. U47296) corresponded to the coding regions 153–173 relative to the first nucleotide of the start codon; siRNAs targeting RL (Acc. No. AF025846) corresponded to region 119–139 after the start codon. The siRNA sequence targeting lamin A/C (Acc. No. X03444) was from position 608–630 relative to the start codon; lamin B1 (Acc. No. NM_005573) siRNA was from position 672–694; NuMA (Acc. No. Z11583) siRNA from position 3,988–4,010, and vimentin (Acc. No. NM_003380) from position 346–368 relative to the start codon. Longer RNAs were transcribed with T7 RNA polymerase from polymerase chain reaction (PCR) products, followed by gel purification. The 49- and 484-bp GL2 or GL3 dsRNAs corresponded to positions 113–161 and 113–596, respectively, relative to the start of translation; the 50- and 501-bp RL dsRNAs corresponded to position 118–167 and 118–618, respectively. PCR templates for dsRNA synthesis targeting humanized GFP (hG) were amplified from pAD3 (ref. 21), whereby 50- and 501-bp hG dsRNA corresponded to positions 121–170 and 121–621, respectively, to the start codon.

For annealing of siRNAs, 20 μ M single strands were incubated in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 min at 90 °C followed by 1 h at 37 °C. The 37 °C incubation step was extended overnight for the 50- and 500-bp dsRNAs, and these annealing reactions were performed at 8.4 μ M and 0.84 μ M strand concentrations, respectively.

Cell culture

S2 cells were propagated in Schneider's *Drosophila* medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS) 100 units ml⁻¹ penicillin, and 100 μ g ml⁻¹ streptomycin at 25 °C. 293, NIH/3T3, HeLa S3, HeLa S6, COS-7 cells were grown at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 100 units ml⁻¹ penicillin, and 100 μ g ml⁻¹ streptomycin. Cells were regularly passaged to maintain exponential growth. Twenty-four h before transfection at 50–80% confluency, mammalian cells were trypsinized and diluted 1:5 with fresh medium without antibiotics (1–3 $\times$ 10⁵ cells ml⁻¹) and transferred to 24-well plates (500 μ l per well). S2 cells were not trypsinized before splitting. Co-transfection of reporter plasmids and siRNAs was carried out with Lipofectamine 2000 (Life Technologies) as described by the manufacturer for adherent cell lines. Per well, 1.0 μ g pGL2-Control (Promega) or pGL3-Control (Promega), 0.1 μ g pRL-TK (Promega), and 0.21 μ g siRNA duplex or dsRNA, formulated into liposomes, were applied; the final volume was 600 μ l per well. Cells were incubated 20 h after transfection and appeared healthy thereafter. Luciferase expression was subsequently monitored with the Dual luciferase assay (Promega). Transfection efficiencies were determined by fluorescence microscopy for mammalian cells lines after co-transfection of 1.1 μ g hGFP-encoding pAD3 (ref. 21) and 0.21 μ g inverted GL2 siRNA, and were 70–90%. Reporter plasmids were amplified in XL-1 Blue (Stratagene) and purified using the Qiagen EndoFree Maxi Plasmid Kit.

Transfection of siRNAs for targeting endogenous genes was carried out using Oligofectamine (Life Technologies) and 0.84 μ g siRNA duplex per well, but it was recently found that as little as 0.01 μ g siRNAs per well are sufficient to mediate silencing. HeLa S6 cells were transfected one to three times in approximately 15 h intervals and were assayed 40 to 45 h after the first transfection. It appears, however, that a single transfection is as efficient as multiple transfections. Transfection efficiencies as determined by immunofluorescence of targeted cells were in the range of 90%. Specific silencing of targeted genes was confirmed by at least three independent experiments.

Western blotting and immunofluorescence microscopy

Monoclonal 636 lamin A/C specific antibody²⁸ was used as undiluted hybridoma supernatant for immunofluorescence and 1/100 dilution for western blotting. Affinity-purified polyclonal NuMA protein 705 antibody²⁹ was used at a concentration of 10 μ g ml⁻¹ for

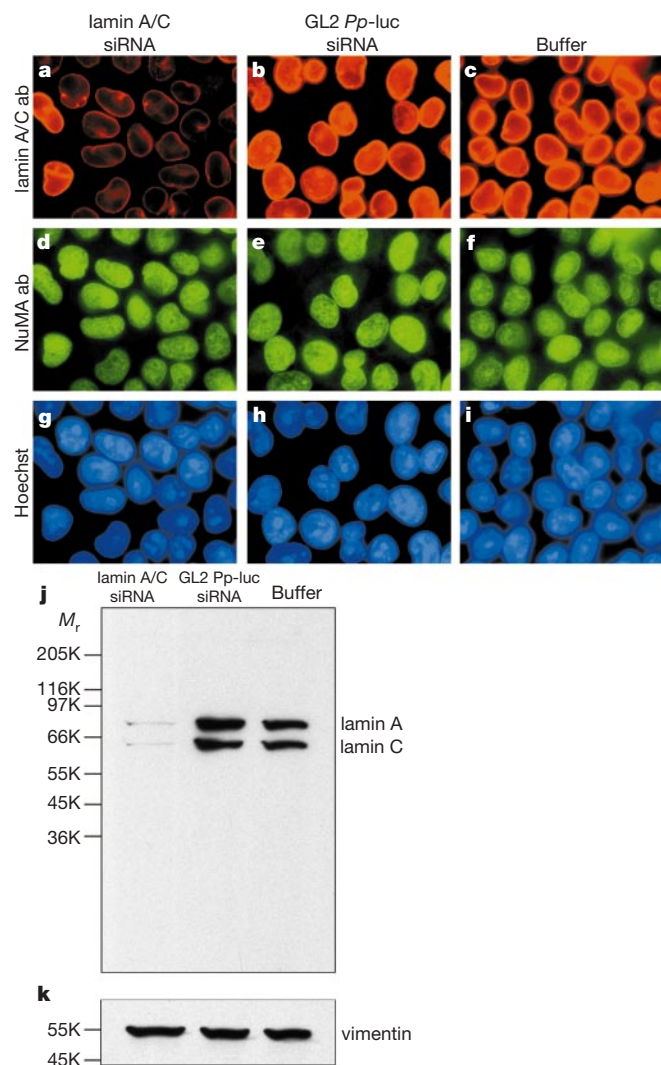


Figure 4 Silencing of nuclear envelope proteins lamin A/C in HeLa cells. Triple fluorescence staining of cells transfected with lamin A/C siRNA duplex (**a, d, g**), with GL2 luciferase siRNA duplex (nonspecific siRNA control) (**b, e, h**), and with buffer only (**c, f, i**). **a–c**, Staining with lamin A/C specific antibody; **d–f**, staining with NuMA-specific antibody; **g–i**, Hoechst staining of nuclear chromatin. Bright fluorescent nuclei in **a** represent untransfected cells. **j, k**, Western blots of transfected cells using lamin A/C- (**j**) or vimentin-specific (**k**) antibodies. The Western blot was stripped and re-probed with vimentin antibody to check for equal loading of total protein.

immunofluorescence. Monoclonal V9 vimentin-specific antibody was used at 1/2,000 dilution. For western blotting, transfected cells grown in 24-well plates were trypsinized and harvested in SDS sample buffer. Equal amounts of total protein were separated on 12.5% polyacrylamide gels and transferred to nitrocellulose. Standard immunostaining was carried out using ECL enhanced chemiluminescence technique (Amersham Pharmacia).

For immunofluorescence, transfected cells grown on glass coverslips in 24-well plates were fixed in methanol for 6 min at -10°C . Target gene specific and control primary antibody were added and incubated for 80 min at 37°C . After washing in phosphate buffered saline (PBS), Alexa 488-conjugated anti-rabbit (Molecular Probes) and Cy3-conjugated anti-mouse (Dianova) antibodies were added and incubated for 60 min at 37°C . Finally, cells were stained for 4 min at room temperature with Hoechst 33342 ($1\ \mu\text{M}$ in PBS) and embedded in Mowiol 488 (Hoechst). Pictures were taken using a Zeiss Axiophot camera with a Fluor 40/1.30 oil objective and MetaMorph Imaging Software (Universal Imaging Corporation) with equal exposure times for the specific antibodies.

Received 20 February; accepted 26 April 2001.

1. Fire, A. RNA-triggered gene silencing. *Trends Genet.* **15**, 358–363 (1999).
2. Sharp, P. A. RNA interference 2001. *Genes Dev.* **15**, 485–490 (2001).
3. Hammond, S. M., Caudy, A. A. & Hannon, G. J. Post-transcriptional gene silencing by double-stranded RNA. *Nature Rev. Genet.* **2**, 110–119 (2001).
4. Tuschl, T. RNA interference and small interfering RNAs. *Chem. Biochem.* **2**, 239–245 (2001).
5. Hamilton, A. J. & Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952 (1999).
6. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296 (2000).
7. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33 (2000).
8. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
9. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21 and 22 nt RNAs. *Genes Dev.* **15**, 188–200 (2001).
10. Caplen, N. J., Fleenor, J., Fire, A. & Morgan, R. A. dsRNA-mediated gene silencing in cultured *Drosophila* cells: a tissue culture model for the analysis of RNA interference. *Gene* **252**, 95–105 (2000).
11. Clemens, J. C. et al. Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl Acad. Sci. USA* **97**, 6499–6503 (2000).
12. Ui-Tei, K., Zennaro, S., Miyata, Y. & Saigo, K. Sensitive assay of RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Lett.* **479**, 79–82 (2000).
13. Wianny, F. & Zernicka-Goetz, M. Specific interference with gene function by double-stranded RNA in early mouse development. *Nature Cell Biol.* **2**, 70–75 (2000).
14. Svoboda, P., Stein, P., Hayashi, H. & Schultz, R. M. Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127**, 4147–4156 (2000).
15. Bahramian, M. B. & Zarbl, H. Transcriptional and posttranscriptional silencing of rodent alpha(I) collagen by a homologous transcriptionally self-silenced transgene. *Mol. Cell Biol.* **19**, 274–283 (1999).
16. Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H. & Schreiber, R. D. How cells respond to interferons. *Annu. Rev. Biochem.* **67**, 227–264 (1998).
17. Manche, L., Green, S. R., Schmidt, C. & Mathews, M. B. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol. Cell Biol.* **12**, 5238–5248 (1992).
18. Minks, M. A., West, D. K., Benvin, S. & Baglioni, C. Structural requirements of double-stranded RNA for the activation of 2',5'-oligoadenylate polymerase and protein kinase of interferon-treated HeLa cells. *J. Biol. Chem.* **254**, 10180–10183 (1979).
19. Clemens, M. & Williams, B. Inhibition of cell-free protein synthesis by pppA^{2'}p^{5'}A^{2'}p^{5'}A: a novel oligonucleotide synthesized by interferon-treated L cell extracts. *Cell* **13**, 565–572 (1978).
20. Macejak, D. G. et al. Inhibition of hepatitis C virus (HCV)-RNA-dependent translation and replication of a chimeric HCV poliovirus using synthetic stabilized ribozymes. *Hepatology* **31**, 769–776 (2000).
21. Kehlenbach, R. H., Dickmanns, A. & Gerace, L. Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT. *In vitro. J. Cell Biol.* **141**, 863–874 (1998).
22. Kreis, T. & Vale, R. *Guidebook to the Cytoskeletal and Motor Proteins*, Parts 2b and 3a (Oxford Univ. Press, Oxford, 1999).
23. Sullivan, T. et al. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J. Cell Biol.* **147**, 913–920 (1999).
24. Wassenegger, M. RNA-directed DNA methylation. *Plant Mol. Biol.* **43**, 203–220 (2000).
25. Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. & Matzke, J. M. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**, 5194–5201 (2000).
26. Wang, M.-B., Wesley, S. V., Finnegan, E. J., Smith, N. A. & Waterhouse, P. M. Replicating satellite RNA induces sequence-specific DNA methylation and truncated transcripts in plants. *RNA* **7**, 16–28 (2001).
27. Razin, A. CpG methylation, chromatin structure and gene silencing—a three-way connection. *EMBO J.* **17**, 4905–4908 (1998).
28. Röber, R. A., Gieseler, R. K., Peters, J. H., Weber, K. & Osborn, M. Induction of nuclear lamins A/C in macrophages in *in vitro* cultures of rat bone marrow precursor cells and human blood monocytes, and in macrophages elicited *in vivo* by thioglycollate stimulation. *Exp. Cell Res.* **190**, 185–194 (1990).
29. Harborth, J., Wang, J., Gueth-Hallonet, C., Weber, K. & Osborn, M. Self assembly of NuMA: multiarm oligomers as structural units of a nuclear lattice. *EMBO J.* **18**, 1689–1700 (1999).
30. Parrish, S., Fleenor, J., Xu, S., Mello, C. & Fire, A. Functional anatomy of a dsRNA trigger: Differential requirement for the two trigger strands in RNA interference. *Mol. Cell* **6**, 1077–1087 (2000).

Acknowledgements

We thank J. Martinez, J. Ludwig and D. Bartel for comments on the manuscript; L. Fredel for help with image processing; H.-J. Dehne for technical assistance; F. Döring, R. Nehring,

D. Ingelfinger and C. Schneider for supplying cell lines; A. Dickmanns for the gift of the plasmid pAD3; and R. Lührmann for support. This work was funded by a Biofuture grant of the Bundesministerium für Bildung und Forschung.

Correspondence and requests for materials should be addressed to T.T. (e-mail: ttuschl@mpibpc.gwdg.de).

Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide

Norbert Polacek, Marne Gaynor, Aymen Yassin & Alexander S. Mankin

Center for Pharmaceutical Biotechnology (MC 870), University of Illinois, 900 South Ashland Avenue, Chicago, Illinois 60607, USA

Peptide bond formation is the principal reaction of protein synthesis. It takes place in the peptidyl transferase centre of the large (50S) ribosomal subunit. In the course of the reaction, the polypeptide is transferred from peptidyl transfer RNA to the α -amino group of amino acyl-tRNA. The crystallographic structure of the 50S subunit showed no proteins within 18 Å from the active site, revealing peptidyl transferase as an RNA enzyme¹. Reported unique structural and biochemical features of the universally conserved adenine residue A2451 in 23S ribosomal RNA (*Escherichia coli* numbering) led to the proposal of a mechanism of rRNA catalysis that implicates this nucleotide as the principal catalytic residue^{2,3}. *In vitro* genetics allowed us to test the importance of A2451 for the overall rate of peptide bond formation. Here we report that large ribosomal subunits with mutated A2451 showed significant peptidyl transferase activity in several independent assays. Mutations at another nucleotide, G2447, which is essential to render catalytic properties to A2451 (refs 2, 3), also did not dramatically change the transpeptidation activity. As alterations of the putative catalytic residues do not severely affect the rate of peptidyl transfer the ribosome apparently promotes transpeptidation not through chemical catalysis, but by properly positioning the substrates of protein synthesis.

The proposed role of A2451 in the peptidyl transfer reaction is

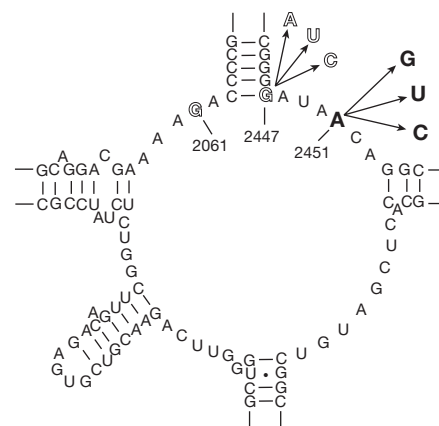


Figure 1 The secondary structure of the central loop of domain V of *T. aquaticus* 23S rRNA. Position A2451 (*E. coli* 23S rRNA numeration), the principal catalytic nucleotide in the proposed general acid–base catalytic mechanism of peptide bond formation^{2,3}, is shown in bold. Its tertiary interaction partners, guanine residues 2061 and 2447, suggested to be essential for rendering catalytic properties to A2451, are outlined. Arrows indicate the mutations engineered in 23S rRNA.

consistent with the results of experiments that implicate this nucleotide in interaction with peptidyl transferase substrates^{4–6}; however, some biochemical and genetic data seem to be in conflict with the proposed catalytic mechanism^{7,8}. Mutations at A2451 (Fig. 1) are dominantly lethal in *E. coli*³ making it difficult to ascertain directly the functional role of this critical nucleotide *in vivo*. Therefore, to investigate the importance of A2451 for peptide bond formation, we used an *in vitro* genetics approach⁹. We engineered three possible nucleotide substitutions at A2451 in the cloned *Thermus aquaticus* 23S rRNA gene. Mutant rRNAs were assembled with ribosomal proteins and 5S rRNA into large ribosomal subunits, and the activities of the mutant subunits were tested in one of the most commonly used peptidyl transferase assays, known as the 'fragment reaction'¹⁰ (assay I; Fig. 2a). In the fragment reaction, the large ribosomal subunit catalyses the transfer of a peptidyl analogue, formyl-[³⁵S]methionyl, from formyl-methionyl-tRNA (fMet-tRNA) to puromycin, a structural analogue of the 3'-end of amino acyl-tRNA. Notably, all three mutants containing base changes at the proposed principal catalytic residue A2451 showed significant peptidyl transferase activity (Fig. 2a). When compared at the end of the linear range of the reaction (at 5 min), the mutant subunits showed activities ranging from 2% (2451G) to 44% (2451U) compared with the reconstituted wild-type 50S subunits (Table 1).

To confirm this finding, we investigated the activities of mutant ribosomal subunits in a substantially different assay. In contrast to the fragment reaction, where the donor substrate (fMet-tRNA) was present in limiting amounts and the acceptor substrate (puromycin) was in excess, in assay II, the relative concentrations of substrates (*N*-acetyl-Phe-tRNA as the donor and 5' [³²P]pCpCp-

puromycin (CC-puromycin) as the acceptor) were reversed. Again all three mutants showed markedly high activity, ranging from 7% (2451U) to 62% (2451G) compared with the wild-type subunits (Fig. 3c; and Table 1). The relative activities of the mutants were switched in these two assays: the 2451G mutant ('weakest' in the fragment reaction) showed the highest activity in assay II; the 2451U mutant ('best' in the fragment reaction) was the least active in assay II. The variation in the relative activities of the mutants in the two assays may reflect differences in interaction with the substrates (puromycin versus CC-puromycin and fMet-tRNA versus *N*-acetyl-Phe-tRNA).

In assays I and II, the peptidyl transfer is catalysed by the large ribosomal subunits in the presence of 33% methanol, which is required to improve the binding of the peptidyl-tRNA¹⁰. It was important to test how mutant subunits would perform under more physiological conditions in the absence of methanol. In the following two assays (III and IV), reconstituted 50S subunits were re-associated with *T. aquaticus* 30S ribosomal subunits to form 70S ribosomes. We then tested the peptidyl transferase activity in the presence of messenger RNA. Assay III employed the same substrates that were used in assay II, and poly(U) was used as mRNA (Fig. 2b). The relative activities of the mutants in the methanol-free assay III were comparable to those in assay II (Table 1). In assay IV, we used a principally new method of measuring peptidyl transferase activity. The method is based on the scintillation proximity assay (SPA)¹¹, which uses streptavidin-coated scintillant-embedded beads. Formyl-[³H]Met-tRNA, which was bound to the ribosome in the presence of synthetic mRNA, served as the peptidyl donor, and biotin-puromycin served as the acceptor. Only after formyl-[³H]methionine is transferred to biotin-puromycin and the resulting compound binds to the streptavidin-coated beads is its radioactivity registered by the embedded scintillant. Both substrates of the peptidyl transferase reaction were present in excess. Under these conditions—where tRNA binding is additionally stimulated by interacting with the 30S subunit—reduction of activity of the mutants compared to the wild type was even less pronounced than in the isolated 50S subunits, ranging from 25 to 73% in assay IV (Fig. 3d; and Table 1).

The actual effect of the mutations on the chemical step of the catalysis can be even smaller than that which follows from our measurements, as mutations at A2451 are expected to affect binding of donor and acceptor substrates to the peptidyl transferase centre. Modification of A2451 with dimethyl sulphate inhibits binding of peptidyl-tRNA to 50S subunits⁵. Thus, one would expect that mutations at this position should interfere with binding of the donor substrates. In agreement with this hypothesis, activity of the 2451C mutant in assay I could be significantly increased (to 78%) by using higher concentrations of fMet-tRNA. Additionally, we observed that activity of the 2451G mutant, which was relatively low in assay I, could be increased roughly 20-fold by using CC-puromycin instead of puromycin (data not shown), showing that the 2451G mutation interferes with binding of puromycin to the 50S subunits. The notion of substrate-binding deficiencies in mutant subunits is further supported by kinetic measurements. In assay I, which was performed under single turn-over conditions, at the time the reaction curves reached their respective plateaux

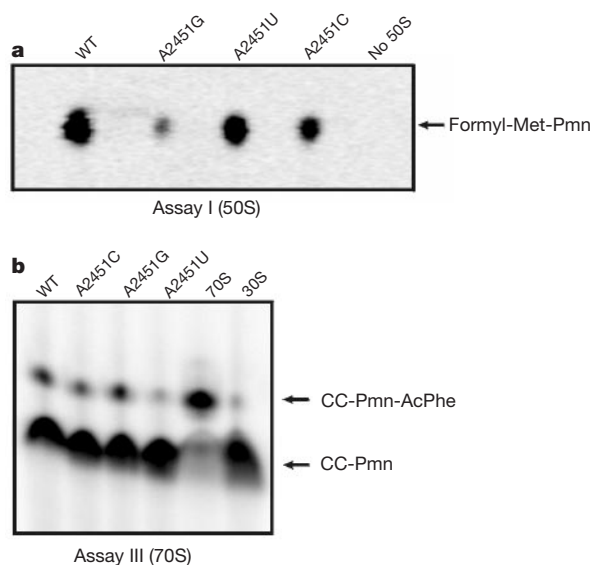


Figure 2 Peptidyl transferase activity of reconstituted large ribosomal subunits harbouring mutations of the proposed catalytic nucleotide A2451. **a**, Assay I. The transfer of formyl-[³⁵S]methionine from tRNA to puromycin was catalysed by 50S subunits in the presence of 33% methanol¹⁰ (see Methods). The reactions were incubated on ice for 30 min, the product, formyl-Met-puromycin (formyl-Met-Pmn), was resolved by paper electrophoresis and visualized by phosphor-imaging. The control (no 50S) contained all the reaction components except for reconstituted 50S subunits. **b**, Assay III. Peptidyl transferase reaction between *N*-acetyl-Phe-tRNA (donor) and 5' [³²P]-CC-puromycin (acceptor) was performed by re-associated 70S ribosomes in the presence of poly(U). The reaction was carried out at 37 °C for 120 min and the reaction product, *N*-acetyl-Phe-CC-puromycin (CC-Pmn-AcPhe), was resolved from CC-puromycin (CC-Pmn) by gel electrophoresis²⁹. The 30S control contained the same amount of 30S subunits as the other samples, but no reconstituted 50S subunits; the observed product spot reflected the activity of minute amounts of native 50S subunits present in the 30S subunit preparation. The 70S control in assay III shows the activity of 30S subunits re-associated with native 50S subunits.

Table 1 Relative activities of *T. aquaticus* large ribosomal subunits

	Assay I (% activity)	Assay II (% activity)	Assay III (% activity)	Assay IV (% activity)
A2451 (WT)	100	100	100	100
A2451U	44 ± 5	7 ± 1	12 ± 5	30 ± 2
A2451C	25 ± 2	31 ± 1	37 ± 13	73 ± 14
A2451G	2 ± 1	62 ± 5	100	25 ± 7

Large ribosomal subunits were reconstituted with wild-type (WT) or mutant 23S rRNA containing mutations at position 2451. Data represent an average of at least two independent experiments for each assay. Incubation times were: assay I, 5 min; assay II, 30 min; assay III, 120 min; assay IV, 240 min.

(5 min) only a portion of fMet-tRNA was converted into the reaction product; less product was formed by the mutants than by wild-type 50S subunits (Fig. 3b). As neither 50S subunits nor reaction substrates are inactivated during this time, only a portion of fMet-tRNA may form reactive complexes with the reconstituted subunits. Assuming that the plateau levels of the reaction curves represent the amount of the reactive complexes at time zero, we found that the observed rate constant (K_{obs}) for wild-type 50S subunit was only two times higher than K_{obs} for the 2451U and 2451G mutants, and four times higher than for the 2451C mutant. Thus, it is possible that reduction of the yield of the peptidyl transfer reaction observed with the mutants comes primarily from impaired binding of the reaction substrates to mutant subunits, rather than from inhibition of chemical catalysis. The observed *in vivo* lethality of mutants at position 2451 in *E. coli*³ may also result from interference with positioning tRNAs on the ribosome.

According to the proposed model², the precise position of A2451 and its N3 group is fixed by a unique hydrogen-bond network with G2061 and G2447. This set of interactions is presumed to be essential for generating an unusually high pK_a at N3 of A2451, thus making it suitable for functioning as a general acid–base in the proposed catalytic mechanism^{2,3}. Mutations of any of these guanines are therefore expected to eliminate the pK_a shift and should severely disturb acid–base catalysis. However, when three possible base changes were introduced at G2447 in *T. aquaticus* 23S rRNA, the activities of reconstituted subunits tested in assays I, II and IV were again high: 18–70% for the adenine mutant, 39–82% for C and 5–48 % for U. Even double mutants harbouring the 2451G mutation, in combination with mutations at the position 2447, yielded subunits with high transpeptidation activities (25–42% in

assay I, for the different mutants; 45–60% in assay II; and 13–22% in assay IV). Our *in vitro* results with the position 2447 mutants are corroborated by *in vivo* data because the G2447U mutation can render cells resistant to a ribosome-targeting antibiotic linezolid^{12,13}, indicating that mutant ribosomes can efficiently support protein synthesis.

Our experiments show that mutations of nucleotides critical for the proposed catalytic mechanism^{2,3} do not abolish peptidyl transferase activity of the ribosome. Mutations of other nucleotide residues in the immediate area of the active site, which could be essential in alternative catalytic models, were also tolerated^{9,14–17}. These findings contrast with results obtained with other ribozymes where alterations of the catalytic residues either completely eliminated catalytic activity or reduce it by many orders of magnitude^{18,19}, indicating that the mode of action of the ribosomal peptidyl transferase differs from that of a number of other ribozymes. The energy required for peptide bond formation is generated during amino acyl-tRNA synthesis coupled with ATP hydrolysis so that peptidyl transferase substrates are delivered to the ribosome in an activated form²⁰. The properly positioned free α -amino group of amino acyl-tRNA is a strong enough nucleophile to spontaneously attack the ester carbonyl group of peptidyl-tRNA. Therefore, the ribosome can potentially promote formation of peptide bonds without significant contribution of chemical catalysis by merely fitting the reaction components in a configuration suitable for the spontaneous reaction^{20,21}. Kinetic considerations suggest that the extent of the acceleration of peptidyl transfer achievable only by substrate orientation can be sufficient to allow for the known rates of protein synthesis²¹. In accordance with this model, the only mutations that have been shown so far to markedly reduce the ribosome's ability to catalyse peptide bond formation *in vitro* are mutations of G2252, a nucleotide that positions peptidyl-tRNA in the ribosomal P-site by base-pairing with the C74 of tRNA^{9,14,22}. The idea that the ribosome catalyses peptidyl transfer primarily by fixing the proper position of the reaction substrates does not eliminate the possible minor contribution of chemical catalysis⁸; however, our results show that it may provide only a small acceleration factor and is apparently not rate limiting.

Although it was not yet possible to demonstrate peptidyl transfer catalysis by isolated rRNA²³, structural and biochemical data leave little doubt that the ribosome is essentially a ribozyme^{2,22,24}, a remnant of the ancient protein-synthesizing machine that could function without the help of proteins²⁵. In the course of evolution, many enzymatic functions thought to be originally performed by RNA enzymes have been taken over by protein catalysts, which provide a broader variety of functional groups that can be used for chemical catalysis. Amino-acid polymerization, however, remained the mission of rRNA. One of the functions that RNA is specifically suited to perform is binding other nucleic acids. The protein synthesis substrates are presented to the ribosome in an already activated form as peptidyl- and amino acyl-tRNAs. Thus, it is possible that, as one of the most ancient enzymes, ribosomal peptidyl transferase evolved simply as a template for binding and proper positioning of its substrates rather than as a chemical catalyst. The apparent limited importance of chemical catalysis explains why this function of rRNA in the ribosome has not been taken over by one of the ribosomal proteins. □

Methods

Preparation of wild-type and mutant 50S subunits

Mutations were engineered at position A2451 (*E. coli* numbering) of the *T. aquaticus* 23S rRNA gene cloned in the pUC18 plasmid²⁶ by PCR mutagenesis using mutagenizing primers 5'-CCATCGATCAACGGATAAAAGTTACCCCGGGGATAXCAGGCT-3' (where X indicates a G, T or C) and a plasmid-specific universal sequencing primer 5'-GTTTCCGACGACGAC-3'. The PCR fragment was cut with *Bsu*151 and *Eco*R1 restriction enzymes and used to replace the corresponding fragment in the wild-type gene. The mutations at position 2,447 were engineered using the same strategy and mutagenizing primers 5'-CCATCGATCAACGGATAAAAGTTACCCCGGGGATAXAC-3', where X

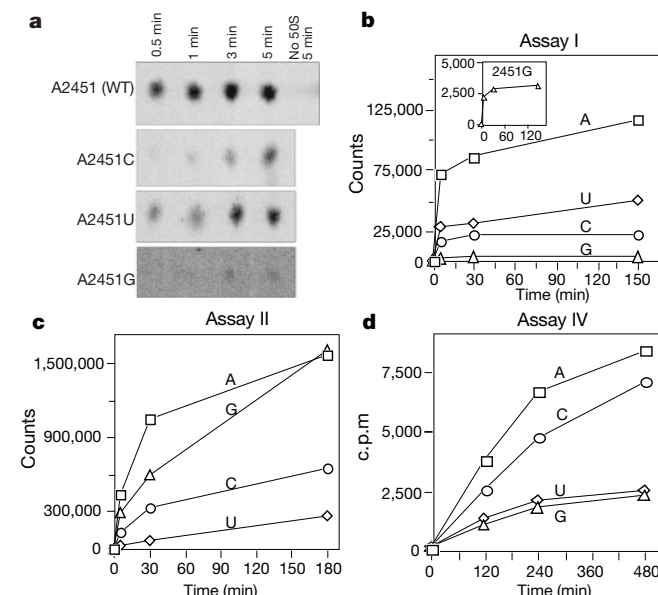


Figure 3 The time course of peptidyl transferase reactions catalysed by reconstituted 50S subunits containing mutations at position 2451. **a**, Phosphor-imager visualization of the appearance of the formyl-[³⁵S]methionine-puromycin product on electrophoregram in assay I. **b–d**, Time course of formation of peptidyl transferase reaction products in assays I (**b**), II (**c**) and IV (**d**). Adenine (wild type, G, C and U) mutants containing corresponding mutations at the position A2451 of 23S rRNA. The inset in **b** represents activity of the A2451G mutant; the time points are the same as on the main graph but the scale of the y-axis was decreased to better reveal the kinetics of product accumulation. The y-axis values in assays I (**b**) and II (**c**) are phosphor-imager arbitrary counts. The radioactivity (c.p.m.) detected by streptavidin-coated scintillant-embedded beads in assay IV (**d**) represents formyl-[³H]methionyl-biotin-puromycin formed as a result of peptidyl transferase reaction (see Methods). Radioactivity values of the corresponding time points of control samples containing only native 30S subunits were subtracted from the experimental values.

indicates A, T or C. Wild-type or mutant 23S rRNA was transcribed *in vitro*, purified, and assembled into 50S subunits as described⁹.

Peptidyl transferase assay I

Reconstituted 50S subunits (21 pmol) (corresponding to 0.5 $\hat{A}_{260}$ of 23S rRNA transcript in reconstitution reaction) were combined with 0.5–1 pmol formyl-[³⁵S]Met-tRNA in 50 μ l of PT buffer (20 mM Tris/HCl pH 8, 20 mM MgCl₂, 0.4 M KCl, 2.5 mM β -mercaptoethanol, 0.1 mM EDTA). After addition of 25 μ l cold methanol, reactions were incubated on ice for 5 min to allow tRNA binding. Peptidyl transfer reaction was initiated by addition of 12.5 nmol puromycin. Reactions were incubated on ice (see figures for specified times), the products were extracted into ethyl acetate and analysed by paper electrophoresis as described²⁷. For the observed rate-constant comparison, experimental points within the first 5 min of incubation were included. The data were fit into the first order exponential decay curve of the enzyme–substrate complex using TableCurve 2D program (SPSS).

Peptidyl transferase assay II

Reconstituted 50S subunits (12.6 pmol), *N*-acetyl-Phe-tRNA²⁸ (13.6 pmol) and 5'-[³²P]CC-puromycin (4 pmol) (Dharmacon Research) were combined in 30 μ l PT buffer containing 33% methanol. Reactions were incubated on ice for the specified time, stopped by addition of 15 μ l 7M urea, and loaded on a 22.5% polyacrylamide/7M urea gel²⁹. Radioactive species were quantified using the Molecular Dynamics phosphor imager.

Peptidyl transferase assay III

Native or assembled *T. aquaticus* 50S subunits (12 pmol) were combined with 30S subunits (12 pmol) in 16 μ l reconstitution buffer^{36,30} (20 mM Tris/Cl pH 7.4, 20 mM MgCl₂, 400 mM NH₄Cl, 6 mM spermidine, 0.4 mM EDTA and 5 mM β -mercaptoethanol) for 1 h at 40 °C. Subsequently, the 70S ribosomes were incubated with 60 μ g poly(U) and 13.6 pmol *N*-acetyl-Phe-tRNA in 30 μ l buffer (final concentrations: 10 mM Tris/Cl pH 7.4, 10 mM HEPES/KOH pH 7.6, 13 mM MgCl₂, 275 mM NH₄Cl, 4 mM spermidine, 0.025 mM spermine, 4.5 mM β -mercaptoethanol, 0.2 mM EDTA). After incubation for 15 min at 37 °C, the reaction was initiated by addition of 4 pmol [³²P]CC-puromycin and allowed to proceed for 2 h at 37 °C. The reaction was stopped and radioactive products were analysed as in assay II.

Peptidyl transferase assay IV

For assay IV we used scintillation proximity assay (SPA) technique¹¹ based on streptavidin-coated scintillant-embedded beads. The radioactivity of the sample is registered only when the ³H source is brought into immediate proximity of the scintillator by binding to the beads. The peptidyl transferase assay used is a modification of the protocol developed by S. Swaney and D. Shinabarger (personal communication). Reconstituted 50S subunits (21 pmol) were reassociated with an equimolar amount of native *T. aquaticus* 30S subunits in 27 μ l reconstitution buffer²⁶ for 1 h at 40 °C. Reconstituted ribosomes were combined with 76 pmol synthetic mRNA AAGGAGAUUAACAUGGGU (Dharmacon Research), 48 pmol formyl-[³H]Met-tRNA (5,900 c.p.m. per pmol) and 60 pmol of a biotin-puromycin derivative (Dharmacon Research) in a final volume of 60 μ l of the adjusted reaction buffer (final concentrations: 20 mM Tris/Cl pH 7.6, 39 mM MgCl₂, 230 mM NH₄Cl, 2.2 mM β -mercaptoethanol, 2.7 mM spermidine and 0.09 mM EDTA). The reaction was performed at 37 °C for 0.5–8 h. The reaction was terminated by addition of 120 μ l stop solution containing 125 mM EDTA, 1 $\times$ PBS and 0.45 μ g streptavidin-coated SPA beads (Amersham Pharmacia Biotech, catalogue number RPNQ0007), transferred to 96-well plates and counted in a 96-well-plate scintillation counter (Wallac). The radioactivity readings of the samples containing only 30S subunits were subtracted for all experimental time points.

Received 26 March; accepted 30 April 2001.

- Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 $\hat{A}$ resolution. *Science* **289**, 905–920 (2000).
- Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
- Muth, G. W., Ortoleva-Donnelly, L. & Strobel, S. A. A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science* **289**, 947–950 (2000).
- Moazed, D. & Noller, H. F. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell* **57**, 585–597 (1989).
- Bocchetta, M., Xiong, L. & Mankin, A. S. 23S rRNA positions essential for tRNA binding in ribosomal functional sites. *Proc. Natl Acad. Sci. USA* **95**, 3525–3530 (1998).
- Steiner, G., Kuechler, E. & Barta, A. Photo-affinity labelling at the peptidyl transferase centre reveals two different positions for the A- and P-sites in domain V of 23S rRNA. *EMBO J.* **7**, 3949–3955 (1988).
- Kearsey, S. E. & Craig, I. W. Altered ribosomal RNA genes in mitochondria from mammalian cells with chloramphenicol resistance. *Nature* **290**, 607–608 (1981).
- Barta, A. *et al.* Mechanism of ribosomal peptide bond formation. *Science* **291**, 203a (2001) (on-line).
- Khaitovich, P., Tenson, T., Kloss, P. & Mankin, A. S. Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with *in vitro*-transcribed rRNA. *Biochemistry* **38**, 1780–1788 (1999).
- Monro, R. E. & Marcker, K. A. Ribosome-catalysed reaction of puromycin with a formylmethionine-containing oligonucleotide. *J. Mol. Biol.* **25**, 347–350 (1967).
- Hart, H. E. & Greenwald, E. B. Scintillation proximity assay (SPA)—a new method of immunoassay. Direct and inhibition mode detection with human albumin and rabbit antihuman albumin. *Mol. Immunol.* **16**, 265–267 (1979).
- Swaney, S. M. *et al.* in *Abstracts of the 38th Interscience Conference on Antimicrobial Agents and*

Chemotherapy C-104 (American Society for Microbiology, Washington DC, 1998).

- Xiong, L. Q. *et al.* Oxazolidinone resistance mutations in 23S rRNA of *Escherichia coli* reveal the central region of domain V as the primary site of drug action. *J. Bacteriol.* **182**, 5325–5331 (2000).
- Green, R. & Noller, H. F. Reconstitution of functional 50S ribosomes from *in vitro* transcripts of *Bacillus stearothermophilus* 23S rRNA. *Biochemistry* **38**, 1772–1779 (1999).
- Green, R., Samaha, R. R. & Noller, H. F. Mutations at nucleotides G2251 and U2585 of 23 S rRNA perturb the peptidyl transferase center of the ribosome. *J. Mol. Biol.* **266**, 40–50 (1997).
- Porse, B. T. & Garrett, R. A. Mapping important nucleotides in the peptidyl transferase centre of 23 S rRNA using a random mutagenesis approach. *J. Mol. Biol.* **249**, 1–10 (1995).
- O'Connor, M. & Dahlberg, A. E. Mutations at U2555, a tRNA-protected base in 23S rRNA, affect translational fidelity. *Proc. Natl Acad. Sci. USA* **90**, 9214–9218 (1993).
- Nakano, S., Chadalavada, D. M. & Bevilacqua, P. C. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **287**, 1493–1497 (2001).
- Yean, S. L., Wuenschell, G., Termini, J. & Lin, R. J. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408**, 881–884 (2001).
- Krayevsky, A. A. & Kukhanova, M. K. The peptidyltransferase center of ribosomes. *Prog. Nucl. Acid Res. Mol. Biol.* **23**, 1–51 (1979).
- Nierhaus, K. H., Schulze, H. & Cooperman, B. S. Molecular mechanisms of the ribosomal peptidyl transferase center. *Biochem. Int.* **1**, 185–192 (1980).
- Samaha, R. R., Green, R. & Noller, H. F. A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. *Nature* **377**, 309–314 (1995).
- Khaitovich, P., Tenson, T., Mankin, A. S. & Green, R. Peptidyl transferase activity catalyzed by protein-free 23S ribosomal RNA remains elusive. *RNA* **5**, 605–608 (1999).
- Zhang, B. & Cech, T. R. Peptide bond formation by *in vitro* selected ribozymes. *Nature* **390**, 96–100 (1997).
- Crick, F. H. C., Brenner, S., Klug, A. & Piecznik, G. A speculation on the origin of protein synthesis. *Orig. Life.* **7**, 389–397 (1976).
- Khaitovich, P. & Mankin, A. S. in *The Ribosome. Structure, Function, Antibiotics and Cellular Interactions* (eds Garrett, R. A. *et al.*) 229–243 (ASM, Washington DC, 2000).
- Green, R. & Noller, H. F. *In vitro* complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for *Escherichia coli* 50S ribosomal subunit assembly and function. *RNA* **2**, 1011–1021 (1996).
- Blaha, G. *et al.* Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. *Methods Enzymol.* **317**, 292–309 (2000).
- Kim, D. F. & Green, R. Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Mol. Cell* **4**, 859–864 (1999).
- Nierhaus, K. H. in *Ribosomes and Protein Synthesis. A Practical Approach* (ed. Spedding, G.) 161–189 (Oxford Univ. Press, Oxford, 1990).

Acknowledgements

We thank M. Gomez and L. Xiong for help with experiments; S. Swaney and D. Shinabarger for providing SPA assay protocols and reagents; and J. Piccirilli, A. Mesecar and A. Neyfakh for helpful discussions. This work was supported by an NIH grant to A.S.M. N.P. was supported by an Erwin Schroedinger fellowship from the Austrian Science Foundation.

Correspondence and requests for materials should be addressed to A.S.M. (e-mail: shura@uic.edu).

Microscopic origins of entropy, heat capacity and the glass transition in proteins

Andrew L. Lee* & A. Joshua Wand

Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Internal motion is central to protein folding¹, to protein stability through the resulting residual entropy², and to protein function^{1,3–7}. Despite its importance, the precise nature of the internal motions of protein macromolecules remains a mystery. Here we report a survey of the temperature dependence of the fast dynamics of methyl-bearing side chains in a calmodulin–peptide complex using site-specific deuterium NMR relaxation methods. The amplitudes of motion had a markedly heterogeneous spectrum

* Present address: Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7360, USA.

and segregated into three largely distinct classes. Other proteins studied at single temperatures tend to segregate similarly. Furthermore, a large variability in the degree of thermal activation of the dynamics in the calmodulin complex indicates a heterogeneous distribution of residual entropy and hence reveals the microscopic origins of heat capacity in proteins. These observations also point to an unexpected explanation for the low-temperature 'glass transition' of proteins. It is this transition that has been ascribed to the creation of protein motional modes that are responsible for biological activity^{5–7}.

The internal motion of protein macromolecules occurs over a wide range of timescales^{1,8}. A variety of spectroscopic methods,

including NMR spectroscopy⁹, have been employed to characterize protein motion in the picosecond to second time regimes. Solution NMR methods offer a very appealing means of studying protein dynamics as they provide comprehensive information resolved at many sites throughout the macromolecule. ¹⁵N-NMR relaxation studies of amide N–H motion generally indicate that the backbone of a stable protein is severely motionally restricted in the picosecond to nanosecond time regime⁹. In contrast, initial views of side-chain dynamics provided by ¹³C-NMR relaxation methods^{10–12} indicated that protein hydrophobic cores are dynamic and, importantly, heterogeneously so. This has significant implications for the presence and internal distribution of residual entropy in proteins^{13,14}. Deuterium NMR relaxation methods¹⁵ allow more efficient investigations than was possible with the ¹³C-NMR relaxation approaches.

We have used deuterium NMR relaxation methods to investigate the changes in conformational entropy of calmodulin upon complexation with a peptide comprising the calmodulin-binding domain of the smooth-muscle myosin light-chain kinase¹⁶. Upon binding to this target domain, calmodulin undergoes a widespread redistribution in side-chain dynamics which is interpreted to correspond to an estimated overall change in conformational entropy of about $-35 \text{ kCal mol}^{-1}$ at 308K. In contrast, the dynamic character of the main chain is largely unaffected by the binding of the target domain.

In an effort to understand the energetics underlying the formation of this high-affinity complex ($K_d \approx 1 \text{ nM}$), we studied the temperature dependence of the fast dynamics of calmodulin. Using methods described in detail elsewhere¹⁶, deuterium NMR relaxation studies of the complex were carried out at 13 temperatures that span 288 to 346K. We probed the dynamics of over 48 methyl-bearing

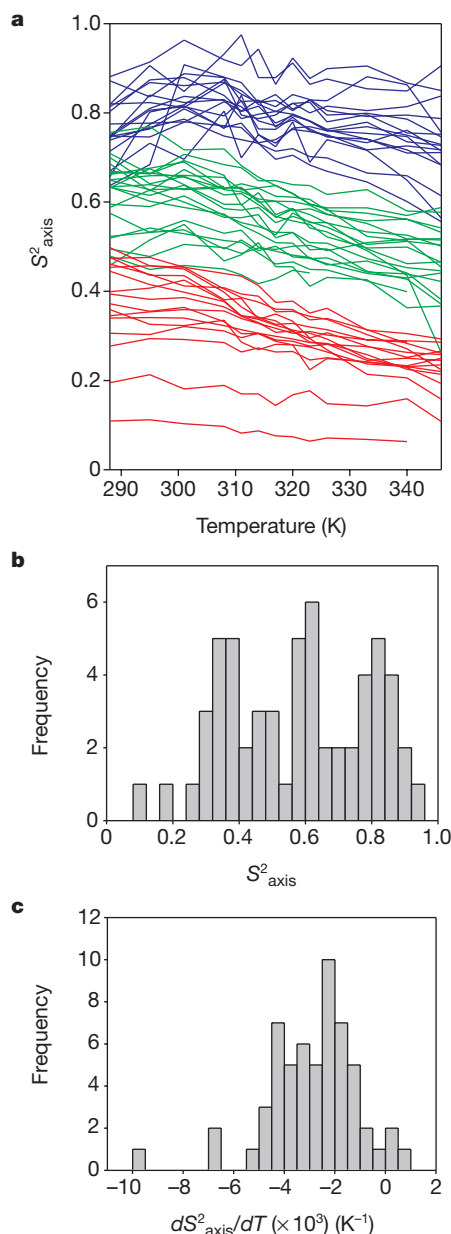


Figure 1 Fast dynamics of methyl-bearing side chains in a calmodulin–peptide complex. **a**, Temperature dependence of S^2_{axis} values of methyl-bearing amino acids of calmodulin in complex with a peptide corresponding to the calmodulin-binding domain of the smooth-muscle myosin light-chain kinase. **b**, Histogram of the distribution of S^2_{axis} values at 308K. **c**, Histogram of the distribution of temperature coefficients of S^2_{axis} values. Data for each site were independently fitted to a linear dependence of the S^2_{axis} on temperature, that is $S^2_{\text{axis}}(T) = 1 - m(T - T_0)$, where m and T_0 are fitted parameters. The 288K data were excluded as S^2_{axis} values at this temperature are significantly less reliable.

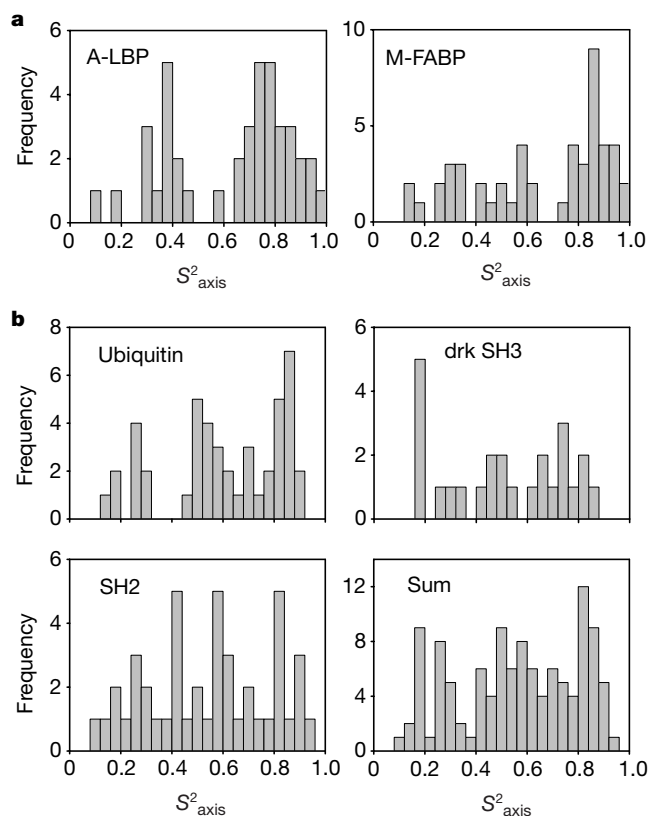


Figure 2 Fast dynamics of methyl-bearing side chains in proteins. **a**, Distribution of S^2_{axis} parameters for human adipocyte (A-LBP) and fatty-acid-binding protein (M-FABP) at 293K. **b**, Distribution of S^2_{axis} values, obtained at 303K, for human ubiquitin¹⁸, the drk SH3 domain²², the phospholipase Cγ-1 SH2 domain²⁰ and their summed distribution.

side chains distributed throughout the calmodulin molecule. The amplitude of motion is expressed as the square of the generalized order parameter of the methyl group symmetry axis (S_{axis}^2) (ref. 17). S_{axis}^2 values can range from unity, corresponding to complete rigidity within the molecular frame, to zero, corresponding to isotropic motion¹⁷. Figure 1 provides a synoptic view of the temperature dependence of the amplitudes of the internal motion undertaken at these sites. This protein is not simply heterogeneously dynamic. Rather, the heterogeneity of the amplitudes of the motions of methyl-bearing side chains of the calmodulin molecule is distributed into three largely distinct bands. The classes of methyl-group dynamics are centred on S_{axis}^2 values of about 0.3, 0.6 and 0.8 at room temperature. With one clear but trivial exception (alanine), the occupancy of a given band is only weakly related to amino-acid type.

To determine whether this trimodal distribution of the amplitudes of fast methyl-bearing side-chain dynamics is a general feature of proteins rather than a feature peculiar to the calmodulin–peptide complex, we examined results from the small number of similar studies of other proteins made at single temperatures^{18–22}. Example histograms of the determined distributions are shown in Fig. 2. Although profiles of individual proteins are limited by the statistics of small numbers, all had non-uniform distributions and indicate that the presence of these three classes of motions in proteins is general.

The banding of internal motion is reminiscent of the hierarchical conformational substate scheme of Frauenfelder *et al.*¹. In this view, the lower band of S_{axis}^2 values, corresponding to the largest angular fluctuations, would arise from motions between various conformational substates in ‘tier 0’. The S_{axis}^2 values of this band correspond to larger-scale motion than the sub-Ångström displacements suggested by studies of crystalline proteins. In the hierarchical scheme, the middle band of S_{axis}^2 values would correspond to interconversion between ‘tier 1’ substates and the top band to interconversion between ‘tier 2’ substates. The fact that the dynamics on all three tiers are represented on the sub-nanosecond timescale was not anticipated and implies that the barriers separating the interconverting states within all three bands are small. Most of the ¹⁵N-derived S_{axis}^2 values lie within the upper half of the top band (not shown). Collectively, these data indicate that the conformational substates sampled on the energy landscape, in the temperature range examined, result primarily from conformational branching at

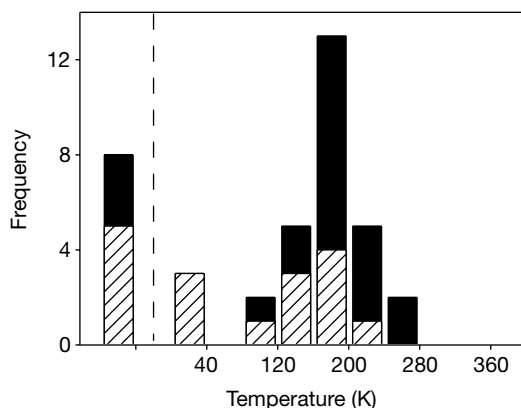


Figure 3 Distribution of the temperatures at which maximal side-chain order is achieved in the calmodulin–peptide complex. The graph shows the distribution of temperatures (T_0) at which individual S_{axis}^2 values extrapolate to unity as described in Fig. 1. Only those sites found in the middle (green) and lower (red) bands of Fig. 1 are included and are represented by solid and hatched bars, respectively. Sites indicated by the bars to the left of the dashed line have sufficiently small temperature dependencies to give negative T_0 values, indicating that they must have significant nonlinear behaviour at lower temperatures.

the side-chain level. Furthermore, unlike the predictions of a purely hierarchical model, most sites do not appear to require a transition between different tiers to explain the observed temperature dependence.

The dependence of S_{axis}^2 parameters on temperature is roughly of the first order between 288 and 346 K (Fig. 1a), as would be expected for largely harmonic motion¹³. All data for each individual methyl site were fitted to a linear dependence upon temperature. The resulting distribution of temperature coefficients (dS_{axis}^2/dT) is broadly centred on a value of about 0.0025K^{-1} (Fig. 1c). These coefficients are, in principle, directly related to the change in local residual entropy with temperature and are, therefore, potentially capable of providing insight into the microscopic origins of heat capacity in proteins. Side chains occupying the top band of Fig. 1a have a shallow dependence upon temperature, which reinforces the assignment of this class to ‘zero point’ motions. Individual temperature dependencies have variable fine structure—some are smoothly linear whereas others have nonlinear features. At this time, no definitive structural or other contextual correlates with the detailed features of the observed temperature dependence. A dissection analysis is ongoing and will be reported elsewhere.

Equilibrium fluctuations, such as those sensed by the deuterium NMR relaxation methods used here, are often interpreted in the context of the statistical mechanics of glasses and glass transitions^{1,8}. A fundamental characteristic of such systems is a qualitative change in the underlying dynamics at or slightly above the glass transition temperature (T_g) (ref. 23). This transition is distinct from the glass transition that is often associated with models of kinetic protein folding⁶. X-ray diffraction^{4–6,24,25} and neutron scattering studies^{26–28} indicate that glass transitions associated with the folded structure occur at about 200 K in proteins. The general view seems to be that this transition involves coupling to solvent and influences the entire protein molecule²⁸. The results presented here indicate an alternative contribution to the apparent glass transition.

Figure 3 shows the distribution of temperatures at which maximal order (T_0) will be reached at individual sites classified into the middle and lower bands of motional amplitude, assuming a linear

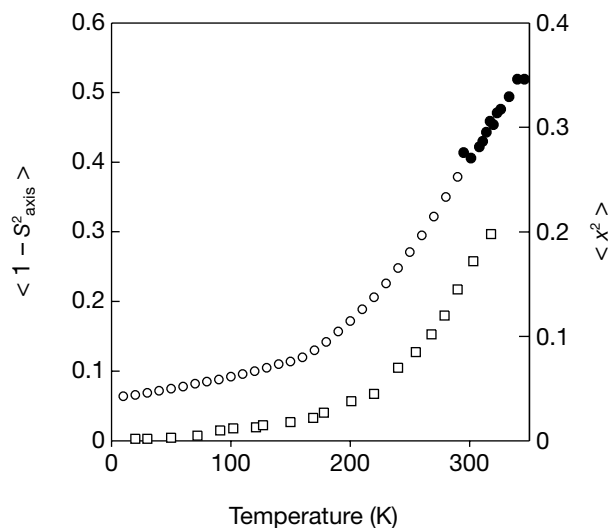


Figure 4 Temperature dependence of the average S_{axis}^2 values of methyl-bearing side chains of complexed calmodulin. $\langle S_{\text{axis}}^2 \rangle$ is the normalized sum of individual S_{axis}^2 values. Black circles correspond to the $\langle S_{\text{axis}}^2 \rangle$ values measured at temperatures above 288 K (see Fig. 1a). White circles correspond to the predicted $\langle S_{\text{axis}}^2 \rangle$ values. At temperatures below a site's T_0 , the corresponding S_{axis}^2 value was set to unity. At temperatures above a site's T_0 , the corresponding S_{axis}^2 value was calculated as described in Fig. 3. The resulting $1 - \langle S_{\text{axis}}^2 \rangle$ curve shown above follows closely the form of the temperature dependence of the mean square atomic displacement ($\langle x^2 \rangle$) obtained from neutron scattering studies of myoglobin (squares)²⁸.

dependence of the S_{axis}^2 parameter on temperature. Most of these sites have T_0 values broadly centred at about 180K, near the temperature often associated with the protein glass transition. An interesting result is obtained if one averages the site-resolved information, obtained here using NMR relaxation methods, to construct an analogue of the average mean squared atomic displacements obtained from neutron scattering studies. Recognizing that the upper limit of the generalized order parameter is unity, one obtains the temperature dependence of $\langle 1 - S_{\text{axis}}^2 \rangle$ shown in Fig. 4. The curve is similar to the temperature dependence of the average mean squared atomic displacements obtained from neutron scattering profiles of several proteins^{26–28}. Typically, such profiles are interpreted to indicate a general transition in the protein–solvent system²⁸. The present result points to a different interpretation. It is the freezing out of the middle band of motion that produces the apparent dynamical transition at about 200K. Furthermore, the upper band, corresponding to the lowest-amplitude motion at room temperature, tends to its limiting value at temperatures well above T_g . The behaviour of the lower band, corresponding to the highest-amplitude motion at room temperature, as the temperature approaches T_g , is less certain. Some members of this class of motion have temperature dependencies that require significant nonlinear behaviour at lower temperatures to converge to maximum order at about 200K (Fig. 3).

We have reported comprehensive site-resolved experimental view dynamics of a protein in solution as a function of temperature. NMR relaxation studies of a calmodulin–peptide complex conducted in solution indicate the presence of three classes of motion which occur on a sub-nanosecond timescale. Similar data for other proteins obtained at single temperatures indicate that this may be a general property of protein molecules. The temperature dependence of the dynamics of the calmodulin complex shows large variations, implying a significant and heterogeneous distribution of residual entropy and, through the associated temperature dependence, reveals local contributions to the heat capacity of the protein. Surprisingly, a simple interpretation of this temperature dependence, at temperatures far above the apparent T_g of proteins, is predictive of the glass transition. This indicates that thermal activation of these motional modes is sufficient to explain the temperature dependence observed by neutron scattering and X-ray diffraction methods without the introduction of a global glass transition model. □

Received 4 January; accepted 3 April 2001.

1. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscape and motions of proteins. *Science* **254**, 1598–1603 (1991).
2. Karplus, M., Ichiye, T. & Pettitt, B. M. Configurational entropy of native proteins. *Biophys. J.* **52**, 1083–1088 (1987).
3. Frauenfelder, H. & Wolynes, P. G. Rate theories and puzzles of hemeprotein kinetics. *Science* **229**, 337–345 (1985).
4. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. Dynamics of ligand binding to myoglobin. *Biochemistry* **14**, 5355–5373 (1975).

5. Rasmussen, B. F., Stock, A. M., Ringe, D. & Petsko, G. A. Crystalline ribonuclease A loses function below the dynamical transition at 220K. *Nature* **357**, 423–424 (1992).
6. Zavodszky, P., Kardos, J., Svingor, A. & Petsko, G. A. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl Acad. Sci. USA* **95**, 7406–7411 (1998).
7. Feher, V. A. & Cavanagh, J. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature* **400**, 289–293 (1999).
8. Onuchic, J. N., Luthey-schulten, Z. & Wolynes, P. G. Theory of protein folding: The energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
9. Palmer, A. G. III. Probing molecular motion by NMR. *Curr. Opin. Struct. Biol.* **7**, 732–737 (1997).
10. Nicholson, L. K. *et al.* Dynamics of methyl groups in proteins as studied by proton-detected ¹³C-NMR spectroscopy. Application to the leucine residues of staphylococcal nuclease. *Biochemistry* **31**, 5253–5263 (1992).
11. Wand, A. J., Urbauer, J. L., McEvoy, R. P. & Bieber, R. J. Internal dynamics of human ubiquitin revealed by ¹³C-relaxation studies of randomly fractionally labeled protein. *Biochemistry* **35**, 6116–6125 (1996).
12. Le Master, D. M. NMR relaxation order parameter analysis of the dynamics of protein side chains. *J. Am. Chem. Soc.* **121**, 1726–1742 (1999).
13. Li, Z., Raychaudhuri, S. & Wand, A. J. Insights into the local residual entropy of proteins provided by NMR relaxation. *Prot. Sci.* **5**, 2647–2650 (1996).
14. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).
15. Muhandiram, D. R., Yamazaki, T., Sykes, B. D. & Kay, L. E. Measurement of H-2 T_1 and $T_{1\rho}$ relaxation-times in uniformly ¹³C-labeled and fractionally H2-labeled proteins in solution. *J. Am. Chem. Soc.* **117**, 11536–11544 (1995).
16. Lee, A. L., Kinnear, S. A. & Wand, A. J. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nature Struct. Biol.* **7**, 72–77 (2000).
17. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559 (1982).
18. Lee, A. L., Flynn, P. F. & Wand, A. J. Comparison of H-2 and C-13 NMR relaxation techniques for the study of protein methyl group dynamics in solution. *J. Am. Chem. Soc.* **121**, 2891–2902 (1999).
19. Mittermaier, A., Kay, L. E. & Forman-Kay, J. D. Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure. *J. Biomol. NMR* **13**, 181–185 (1999).
20. Kay, L. E., Muhandiram, D. R., Farrow, N. A., Aubin, Y. & Forman-Kay, J. D. Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry* **35**, 361–368 (1996).
21. Constantine, K. L. *et al.* Backbone and side chain dynamics of uncomplexed human adipocyte and muscle fatty acid-binding proteins. *Biochemistry* **37**, 7965–7980 (1998).
22. Yang, D., Mittermaier, A., Mok, Y. K. & Kay, L. E. A study of protein side-chain dynamics from new ²H auto-correlation and ¹³C cross-correlation NMR experiments: application to the N-terminal SH3 domain from drk. *J. Mol. Biol.* **276**, 939–954 (1998).
23. Sokolov, A. P. Why the glass transition is still interesting. *Science* **273**, 1675–1676 (1996).
24. Hartmann, H. *et al.* Conformational substates in a protein: structure and dynamics of metmyoglobin at 80K. *Proc. Natl Acad. Sci. USA* **79**, 4967–4971 (1982).
25. Tilton, R. F. Jr, Dewan, J. C. & Petsko, G. A. Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320K. *Biochemistry* **31**, 2469–2481 (1992).
26. Doster, W., Cusack, S. & Perry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **337**, 754–756 (1989).
27. Diehl, M., Doster, W., Petry, W. & Schöber, H. Water-coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering. *Biophys. J.* **73**, 2726–2732 (1997).
28. Réat, V. *et al.* Solvent dependence of dynamic transitions in protein solutions. *Proc. Natl Acad. Sci. USA* **97**, 9961–9966 (2000).

Acknowledgements

We thank S. W. Englander and K. Sharp for discussions. This work was supported by a grant from the NIH.

Correspondence and requests for materials should be addressed to A.J.W. (e-mail: wand@mail.med.upenn.edu). The relaxation data has been deposited in the BMRB under accession number 4970.

Microscopes under the microscope

Microscopes, cameras, slides and knives.

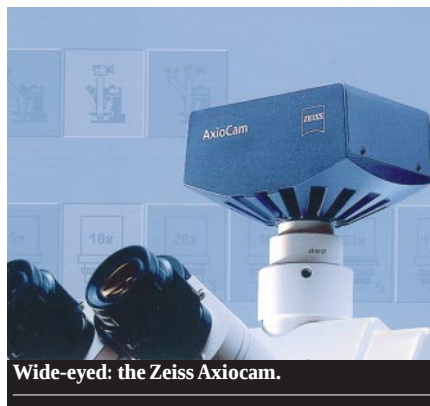
Axiocam

Zeiss

www.zeiss.de

A broader view

The Axiocam ultra-high-resolution camera uses a scanning method that gives wide fields of view with resolutions from $1,000 \times 1,300$ pixels to $3,090 \times 3,900$ pixels at 14-bit colour depth. It can integrate images captured over time, facilitating the acquisition of fluorescently labelled specimens. Peltier cooling helps keep background 'noise' to a minimum, enhancing length of integration and operating range. Axiovision software allows the user to shorten the speed of acquisition and keep file size small by lowering the resolution. The scanning mode can be switched off and the image can also be acquired as an interpolated image. Suggested applications include brightfield capture, phase contrast, DIC, darkfield and fluorescence imaging.



Wide-eyed: the Zeiss Axiocam.

and $\times 80$ magnification, with no loss of resolution, using an intermediate slider. The specimen is not disturbed and microscope stability is maintained. Light collection facilitates observation deep within the specimen, making these new microscopes particularly suitable for sensitive applications such as patch-clamping, gene injection into thick samples and membrane potential dye imaging.

ISIS

Vision Engineering

www.visioneng.com

Far-sighted strategy

Using the ISIS eyepiece increases viewing distance to 38 mm, allowing operators to wear prescription or safety glasses and avoid physical contact with the unit. Unlike traditional microscope eyepieces, the ISIS does not depend on precise alignment of the eyes with the centre of each eyepiece, so movement of the pupils does not result in losses in the field of vision. The device can be fitted into binocular stereo and research microscopes, including those manufactured by Leica, Nikon, Olympus, Wild and Zeiss.



BX51WI, BX61WI

Olympus

www.olympus.com

The objective is the same

Olympus has introduced two new fixed-stage upright microscopes designed for the observation of living cell specimens. Both instruments feature a $\times 20$ objective with an NA of 0.95, allowing the user to view the specimen at $\times 7$

DC200

Leica Microsystems

www.leica-microsystems.com

Fast forward

The DC 200 digital microscope camera is equipped with real-time image preview with nine frame per second colour and 17 frame per second black-and-white imaging. The high preview speed allows users to focus and magnify directly from a computer monitor, reducing the need to use the microscope eyepieces. A $2/3$ -inch chip provides $1,280 \times 1,024$ resolution for presentation and publication of data, enhanced by 30-bit colour depth in both preview and captured images.

SMZ1500

Nikon

www.nikonusa.com

Light up your life science

Nikon's new stereo microscope has a zoom ratio covering a range from $\times 0.75$ to $\times 11.25$, allowing life scientists to see and photograph any specimen, from macro views to high-magnification micro visualization. No auxiliary lenses are necessary. Depending on the eyepieces and the objective selected, the SMZ1500 produces total magnifications from $\times 3.75$ to $\times 540$. Optional 'oblique coherent contrast' is also available. This new illumination technique provides enhanced surface imaging and is suggested for cell

harvesting, *in vitro* fertilization and transgenics applications.

Spot RT

Diagnostic Instruments

www.diaginc.com

Spot the difference

Designed specifically for microscopy, the Spot RT Monochrome digital camera, with the new Kodak KAI-2000 CCD, provides a 19 f.p.s. live image window for focusing and scanning as well as 12-bit image captures. The research-grade CCD produces a resolution of $1,600 \times 1,200$, almost two megapixels of non-interpolated data. Cooling to -37°C ensures low background noise for fluorescence and darkfield work. A broad-band coated UBK7 window is standard and fused silica or infrared blocking windows for imaging live specimens are available as options. The camera is compatible with both Macs and PCs.

Volocity

Improvision

www.improvision.com

See its true colours

This four-dimensional rendering system is designed for biomedical imaging. It uses advanced algorithms to provide high-speed, interactive rendering of time-resolved colour three-dimensional volumes. The system allows the user to visualize a three-dimensional object and then observe and interact with it over time, thus allowing dynamic visualization of both the structure and the function of cellular structures.

ImmunoGold Newsletter

Electron Microscopy Sciences

www.emsdiasum.com

Keep up with new agents

The EMS newsletter features ImmunoGold reagents and accessories manufactured by Aurion. It contains a listing of all available immuno reagents, a section of frequently asked technical questions, some sample protocols, a selection of micrographs and an overview of custom labelling services. Also available via e-mail from sgkcc@aol.com.

Near-field scanning optical microscope

Veeco Metrology

www.di.com

Fully integrated fluorescence microscopy

The Veeco Metrology Group, part of Digital Instruments, has introduced a new SEM system based on the NanoScope AFM controller and the Zeiss Axiovert 135 inverted optical microscope. High-efficiency optics and integrated detectors allow simple, high-resolution transmission and fluorescence maps, in addition

Olympus: maintaining a stable view

tion to topography on the nanometre scale. The system is designed to incorporate an acousto-optic laser module for simultaneous mixing of multiple wavelengths. Integration of the fibre scanner unit in the condenser lens allows both conventional far-field microscopy and near-field microscopy of the same sample. Suggested applications include high-resolution mapping of fluorescent labels, single-molecule fluorescence and FRET, liquid crystal polarization and magneto-optics.

Nano-Bio3

Mad City Labs www.madcitylabs.com

Triple axis nanopositioning system

This system from Mad City Labs (they are from Madison) offers sub-nanometre accuracy under closed loop control and permits triple axis absolute imaging of cells, single molecules and other nanostructures. Its low-profile design allows it to be easily integrated into existing inverted microscopes and the Digital Instruments BioScope. A generous through-hole is provided to permit the admission of probes or objective lenses without compromising range of motion. The system features parallel, uncoupled motion in three axes, with 100 μm of travel in x and y axes and 10–50 μm of travel in the z axis. It has integrated piezoresistive sensors for position measurement. Available in either aluminium or, if thermal stability is a priority, invar, the Nano-Bio3 can also be ordered in two- and five-axis versions.

Cryo-P

Diatome www.emsdiasum.com

Diamond knife

This knife for cryo sectioning facilitates sliding of dry sections on the diamond surface and pick-up from the platform. Semi-thin and ultra-thin sections can be manipulated with the same knife, and the platform is said to reduce structural damage to tissues and cells. It is located on the back of the knife, towards the user, enabling easy retrieval from a dark epoxy surface. The knife's surface provides an area of high visibility for sections normally difficult to see on lower-contrast coatings.

DMC2

Polaroid www.polaroid.co.uk/dmc

A print-friendly digital microscope camera

Polaroid's DMC2 high-resolution digital microscope camera provides output resolution of $1,600 \times 1,200$ with 14-bit per channel linear RGB images. Features include an electronic shutter, exposure times up to 8 seconds, useful for low-light applications and fluorescence imaging, rapid-fire 'burst mode' image capture, resizable preview window and an automatic exposure facility. A full-colour preview rate of up to 12 frames per second optimizes image capture. Colour Stacker software enables the user to combine up to three images

into one and is compatible with any microscope having a C-mount adaptable photoport. Images can be saved in 8- or 16-bit TIFF format or other image file types available with Adobe Photoshop for Mac or TWAIN for Windows.

BioCoat coverslips

BD Biosciences www.bd.com

Slip into a bargain

BioCoat coverslips with extracellular matrix proteins and attachment factors are available from BD Biosciences at reduced prices and in more convenient packaging. The coverslips are recommended for applications requiring optically clear, thin glass bottoms, including immunofluorescence microscopy, fluorescence imaging of live cells, confocal microscopy, high-resolution, inverted and phase-contrast microscopy.

X-Cite

EFOS www.efos.com

Snap out of it with this lamp

The X-Cite lamp incorporates a pre-aligned snap-in/snap-out lamp with automatic monitoring of lamp temperature to prevent damaging lamp restrikes. The shutter system has a foot-operated control to free the user's hands for navigation and control of the microscope. The spectral range spans 250–600 nm, limiting undesired sample heating and making the system particularly suitable for live cell work.

Radiance 2100

Bio-Rad www.bio-rad.com

Confocal microscopy to build on

The Radiance 2100 confocal microscope can be progressively upgraded from single-channel through to multi-channel, multi-photon as required. Confocal microscopy enables the acquisition of optical section images through a fluorescent biological sample so that only light in focus is collected. The Radiance 2100's alignment-free, laser point scanning systems is designed for low maintenance. Each system offers a choice of emission filters combining high spectral resolution with 100% repro-



BioCoat coverslips, now on offer.

ducibility and independence of confocal aperture size or wavelength. The system is compatible with a wide range of upright and inverted microscopes from Nikon, Zeiss and Olympus.

Vivid Filters for Microscopy

Omega Optical www.omegafilters.com

Electronic catalogue launched

'Vivid Filters for Microscopy' is a new web-based catalogue describing Omega's filter sets and components for fluorescence microscopy, including filters for GFP, multi-labelling, ratio imaging and other applications. Filter sets are grouped according to design features and application, and are listed in conjunction with information about their component filters, including spectral descriptions.

Filter system

Prior Scientific www.prior.com

Re-inventing the wheel

Smooth, high-speed operation is promised by Prior for its new filter wheel, which can change filters in as little as 55 ms. A choice of wheels is available capable of accepting either ten 25 mm or eight 32 mm diameter filters. Up to two filter wheels and three shutters can be operated from a single controller. The system is fully programmable via a serial port or, for applications without a host computer, an optional filter wheel keypad is available.

These notes are compiled in the Nature office from information provided by the manufacturers.

ADVERTISEMENTS

**BACs • YACs • PACs
cDNAs**

We provide clones and colony membranes for a number of human, mouse and rat libraries as well as a custom screening service for the libraries.

ResGen™
Invitrogen Corporation

U.S. or Canada 800-533-4363
FAX 256-536-9016

Check out our web site at
www.resgen.com

**DIVERSE
SMALL
MOLECULES
READY FOR
SCREENING**

**10K–50K
Compounds
in Plates**

- Next Day Delivery*
- Competitively Priced

*(Some Restrictions Apply)

Chembridge Corporation • 800.964.6143
858.451.7400 • 858.451.7401 FAX • E-mail: chem@chembridge.com

www.chembridge.com

Contacts

Publisher: Fabien Savenay

Editor: Paul Smaglik

Sales Director: Ben Crowe

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: classified@nature.com

Group European Manager:

Leonie Welss (4954)

European Manager:

Nevin Bayoumi (4978)

European Sales Consultant:

Jonathan Bull (4973)

UK/ RoW/ Ireland:

Matt Powell (4953), Ben Corp (4974),

Andy Douglas (4975)

Holland/ Italy: Nevin Bayoumi (4978)

Scandinavia: Sille Opstrup (4994)

Production Manager:

Billie Franklin
To send films and materials use
London address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: classified@nature.com

International

Advertising Coordinator:

Laura Pearson (4977)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Offices

France/ Belgium:

Christine Niox-Chateau

Tel + 33 (0) 1 43 20 16 51

Fax + 33 (0) 1 43 20 51 52

e-mail: c.nioxchateau@nature.com

Spain/ Portugal:

Amanda Wren

Tel + 34 91 3913835

Fax + 34 91 319 0091

e-mail: a.wren@nature.com

Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mails: p.phelan@nature.com

k.turner@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: classified@nature.com

US Sales Director:

Ben Crowe

US Sales Manager:

Peyton Mason

US Advertising Coordinator:

Corissa Salzman

Japan Head Office, Tokyo

Shin-Mitsuke Building (4F),
3-6 Ichigaya Tamachi Shinjuku-Ku,
Tokyo 162

Tel +81 3 3267 8751

Fax +81 3 3267 8746

e-mail: k.cowan@nature.com

Japan Manager:

Kate Cowan

Trend or aberration?

Aneecdotes about scientists leaving universities for industry jobs used to warrant curiosity — or worse, whispers of 'sell-out' from envious colleagues. But the exodus has become so commonplace that it no longer draws attention, beyond speculation over money. Tales about scientists departing industry for academia are more exotic. So when the moves of two prominent scientists from large corporations to university positions come to light more or less at the same time, one must resist the temptation to label this reverse migration a trend. Instead, it may be more prudent to examine the conditions that prompted those moves, and ask if they are aberrations or harbingers of things to come.

Mathematician Andrew Odlyzko and nanotech pioneer Jim Gimzewski both enjoyed the resources of industry for a combined 40-plus years — Odlyzko leading AT&T's mathematics and cryptography department (see *Movers*, *Naturejobs* 17 May 2001) and Gimzewski manipulating atoms at IBM's famed Zurich Research Centre (see *Movers*, this issue). But both eventually found that their interests were broader than their companies' relatively open playing fields could contain. Moving to universities means more control over intellectual property, a chance to interact with hundreds of postdocs and graduate students, and the opportunity to form and dissolve collaborations with colleagues in both sectors.

Still, even with that freedom, the decision was not easy for either scientist. The question is, will more scientists be asked to make it? Both Odlyzko and Gimzewski will run large, well-funded, multidisciplinary centres. Although there are relatively few of these centres, their numbers are growing and it is not unreasonable to speculate that as more are opened, academic scientists will no longer find it curious that their industrial colleagues are joining them.

Paul Smaglik

Naturejobs editor



Contents

MOVERS

Nanotech pioneer leaves IBM for academia; Princeton gets first woman president; and more

[Back page](#)

WWW.NATURE.COM/ NATUREJOBS

Career centre
Information on the scientific
job market

FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS

NANOTECHNOLOGY

Jim Gimzewski has left IBM Zurich Research Laboratories to run his new nanolab in the department of chemistry and biochemistry at the University of California, Los Angeles. He has been with IBM since 1983 and joined the corporation after a University of Zurich postdoc spent studying plasma surface interactions. IBM originally wanted Gimzewski to apply its scanning tunnelling microscopes (STMs) to industrial processes, but the nascent technology was not then fully developed. Instead, he used STMs to study organic molecules, construct tiny electrical contacts and manipulate atoms into new configurations. Although he enjoyed his years at the Swiss lab, Gimzewski notes that the industrial atmosphere gradually provided him with less freedom to pursue those independent lines of inquiry in basic research. "IBM is more focused on applied research these days," he says. "They're much more directed." At the nanolab, which is part of California's \$250-million nanotech initiative, he can pursue his interests as they shift and evolve. Right now, that means exploring the chemical and biochemical aspects of nanotechnology, rather than the physical properties, which was his focus at IBM. He also looks forward to working with scientific teams whose student and postdoc memberships continue to evolve, along with their research interests. He also notes that, as a university researcher, he gets more rights to his intellectual property than he would with a company.

ENERGY

T. J. Glauthier, former deputy secretary and chief operating officer at the US Department of Energy (DOE), has been appointed president and chief executive of the Electricity Innovation Institute. The Palo Alto, California-based non-profit institute, which is affiliated with the Electric Power Research Institute, aims to spur public-private research collaborations in energy. Glauthier served over seven years in executive positions for former US President Bill Clinton, including five years at the White House in the Office of Management and Budget, before moving on to the DOE. Before joining the Clinton administration, Glauthier directed the World Wide Fund for Nature's climate-change campaign. He also worked for over 20 years in management consulting as vice-president of the public policy and management group at Temple, Barker, and Sloane.

ADMINISTRATION

Molecular biologist **Shirley M. Tilghman** has been named this month as the first woman president of Princeton University, and only the second-ever woman president of an Ivy League university. Tilghman, who will take up the position on 15 June, is perhaps known for her policy work as much as for her science. She was a co-author of the US National Academy of Sciences report, *Trends in the Early Careers of Life Scientists*, which analysed the overproduction of life-science PhDs in the United States. The work, which reverberated widely through policy circles, is often referred to as the 'Tilghman Report'. She has also served on national advisory bodies that dealt with embryonic stem-cell research and human cloning. Tilghman, who has been a member of the Princeton faculty since 1986, served as the first director of the multidisciplinary Lewis-Sigler Institute for Integrative Genomics, formed in 1999. She will replace Harold T. Shapiro, who has held the position for some 13 years.

BIOLOGY

Donna Dean will serve as acting director of the newly formed US National Institute of Biomedical Imaging and Bioengineering (NIBIB). The NIBIB, the newest US National Institutes of Health institute, was created by statute and was signed into law and authorized by President George W. Bush's administration last December. Dean has served in the Office of the NIH Director as a senior scientific adviser for the past three years, and played a lead role in implementing the legislative establishment of NIBIB. Dean also served as NIH liaison to the congressionally mandated Commission on the Advancement of Women, Minorities, and Persons with Disabilities in Science, Engineering, and Technology during the past two years. The president has requested \$40.2 million for the new institute for fiscal year 2002, which begins on 1 October.

Husband and wife **Jonathon Howard** and **Karla Neugebauer** are joining the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden next month. Both now work at the University of Washington in Seattle. Howard was a professor of physiology and biophysics and Neugebauer was an assistant professor of neurology. The Dresden institute opened its doors this January.



Jim Gimzewski



T. J. Glauthier



Shirley Tilghman



Donna Dean

CONTACTING US AT MOVERS

Please send any information on new appointments or job moves to: naturejobseditor@naturedc.com.